STRUCTURE AND FUNCTION OF THE G DOMAIN OF PARKINSON'S DISEASE-ASSOCIATED PROTEIN LRRK2

Chunxiang Wu

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the Department of Biochemistry and Molecular Biology,
Indiana University

August 2019

Accepted by the Graduate Faculty of Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee

_____
Quyen Q. Hoang, PhD, Chair

_____
Tatiana M. Foroud, PhD

January 17, 2019

_____
Thomas D. Hurley, PhD

_____
Steven M. Johnson, PhD

_____
Zhong-Yin Zhang, PhD

DEDICATION

This thesis is dedicated to my mother and father, Fang-Ya and Ping, for all the love, support and understanding.

To my dear friends and lab mates for the encouragements, advices and all the fun and happy times.

Lastly, I dedicate this thesis to my mentor Dr. Quyen Hoang, thank you for the guidance, patience and inspiration throughout this project; for providing the warm and friendly lab life; and for all the Thanksgiving meals that make me feel at home.

ACKNOWLEDGEMENT

Firstly, I would like to express my profound appreciation to my advisor, Professor Quyen Q. Hoang, for allowing me to join our lab and supporting me in the study of this exciting, interesting and significant research project. His motivation, enthusiasm, immense knowledge and patience in science deeply inspired and encouraged me to learn more and work harder. I would always remember the first time coming to him and seeing the quote at the door of his office: "Wisdom is knowing what to do next, skill is knowing how to do it, and virtue is doing it." Working together with him, I got trained not only with the scientific skills, but also precise logical thinking and the feeling of responsibility to contribute the understanding of health and disease.

Also, I would like to express my deep appreciation to my thesis committee advisors: Professors Thomas D. Hurley, Zhong-Yin Zhang, Tatiana Foroud and Steven M. Johnson, for their insightful comments and advises that helped me realize overlooked aspects and possibilities, design and modify my research plans with better strategies, and encourage me to overcome the obstacles during my research project.

In addition, I would like to thank Dr. Lan Chen, Dr. Jingwei Meng, Dr. Lifan Zeng, Dr. Mu Wang, Dr. Amber Mosley, Guihong Qi for helping me with assistance and great advises on my experiments in the chemical genomics core and proteomics core facilities. I would also like to thank Dr. Ruslan Sanishvili at the GM/CA beamline of the Advanced Photon Source, for his ingenious ideas and help during the X-ray diffraction data collection at Argonne National Laboratory.

I am thankful to Dr. Millie M. Georgiadis, Dr. Yuichiro Takagi, Dr. Ann C. Kimble-Hill and everyone else in our structure biology group meeting, where I received many

insightful comments and advises while also widened my scientific vision from exciting presentations on various projects. I would like to extend my appreciation to Dr. Mark Goebl and the Biochemistry Office staff for their kindly and timely help and assistance.

I also want to thank my previous lab colleagues: Dr. Jingling Liao, Dr. Wei Wang, who gave me an unforgettable warm and happy start in the lab and trained me with many skills so that I could quickly "jump-in" my research project. I would also thank my current lab colleagues Dr. Yangshin Park, Alexe Engel, Dr. Misook Oh, and Dr. Li Wan for all the fun and happiness in the lab. Also, I want to thank my dear friends: Dr. Liang Wang, Dr. Jie Lan, Dr. Qiujia Chen, Chen Chen and Shijun Zhang, for their encouragements and moral supports that make me feel like at home during my stay in Indianapolis.

Lastly, I would also like to thank my parents, grandparents and all families in China, for supporting and encouraging me to take on this great adventure oversea to study and pursue on exciting scientific researches at Indiana University School of Medicine.

.

Chunxiang Wu

STRUCTURE AND FUNCTION OF THE G DOMAIN OF PARKINSON'S DISEASE-ASSOCIATED PROTEIN LRRK2

Mutations in the gene encoding for leucine rich repeats kinase 2 (LRRK2) are commonly found in Parkinson's disease. Recently, we found that the disease-associated point mutations at residue R1441 in the G domain (ROC) of LRRK2 resulted in perturbation of its GTPase activity. In this study, we compare the biochemical and biophysical properties of the ROC domain of LRRK2 carrying the PD-associated mutations at residue R1441 with those of the wild-type. We found that the disease-associated mutations (R1441C/G/H) showed marked quaternary structure compared to wild-type, in that the latter existed in solution in both monomeric and dimeric conformations dynamically regulated by GDP/GTP binding state, while we detected only monomeric conformation for three disease-associated mutants. To understand the structural basis for this plasticity and the activity reduction in the mutants, we solved a 1.6 Å crystal structure of the wild type ROC that shows a stable dimeric conformation in which the switch motifs and inter-switch regions mediate extensive interactions at the dimer interface. Residue R1441, where PD-associated mutations occur, forms exquisite interactions at the interface, thus suggesting a critical role of this residue in maintaining a dynamic dimer-monomer interconversion and conformational flexibility of the switch motifs. Consistently, substituting R1441 for other arbitrary mutations (R1441K/S/T) lead to similar perturbation of GTPase activity and dimerization defects as observed in the disease-associated mutants. Locking the ROC domain in either dimeric or monomeric conformations by engineered disulfide bond alters the binding affinity to GTP (but not GDP) and significantly reduce

GTPase activity, thus suggesting that the dynamic dimer-monomer interconversion and conformational plasticity are essential for ROC function as a molecular switch modulating the kinase activity of LRRK2.

Quyen Q. Hoang, PhD, Chair

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ARM	Armadillo repeats domain
ANK	Ankyrin repeats domain
BODIPY	Boron-dipyrromethene
COR	C-terminal of ROC domain
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
FP	Fluorescence polarization
GAD	G protein activated by dimerization
GAP	GTPase activating protein
GDI	Guanosine nucleotide dissociation inhibitor
GDP	Guanosine diphosphate
GppCp	Guanosine-5'-[(β,γ)-methyleno]triphosphate
GppNHp	Guanosine-5'-[(β,γ)-imido]triphosphate
GEF	Guanine nucleotide exchange factor
GTP	Guanosine triphosphate
GTPγS	Guanosine 5'-O-[gamma-thio]triphosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LMNG	Lauryl maltose neopentyl glycol
LRR	Leucine rich repeat domain
LRRK2	Leucine rich repeat kinase 2
MALS	Multiple angle light scattering

PCR	Polymerase chain reaction
PD	Parkinson’s disease
PDB	Protein data bank
PEG	Polyethylene glycol
PINK1	PTEN-induced putative kinase
ROC	Ras-like G domain of complex protein
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SNCA	Alpha synuclein
UCHL1	Ubiquitin C-terminal hydrolase
VPS35	Vacuolar protein sorting protein 35

Chapter 1. Function and oligomeric plasticity of the GTPase domain of the Parkinson disease-associated protein LRRK2

1.1 Introduction

A. An overview of Parkinson's disease: Neuronal degeneration and Lewy pathology

The world is facing with an unprecedented problem associated with an aging population in the coming decades, with over 25 % of the population reaching 65 years of age or older within the next 50 years. This would create an unmanageable situation for world health care systems (Petsko 2006).

Parkinson's disease (PD) is one of the most common age-associated diseases. While the incidence of PD is about 1.5 % of population over 65, the risk after the age of 85 increases more than 4-fold (Lees, Hardy et al. 2009, Stoessl, Martin et al. 2011). Currently, there are more than 1 million PD patients in the US; however, due to the rapidly aging population, the number of the PD patients is estimated to rise to 3 million over the next 50 years (Petsko 2006). Parkinson's disease (PD) was first described by the English surgeon James Parkinson in his monograph entitled "An Essay on the Shaking Palsy" in 1817. Thus, it was later named by Jean Martin Charcot as maladies de Parkinson (Parkinson's disease) (Kempster, Hurwitz et al. 2007). The clinical symptoms of PD include resting tremor, bradykinesia (slowness of movement), rigidity, and postural instability (balance problems). These characteristic features are still being used in the diagnosis for PD today.

Although the pathologic mechanism of PD is still largely unknown, the debilitating symptoms of PD are caused by the progressive neuronal degeneration of the dopaminergic

cells in the substantia nigra pars compacta (Figure 1). These neurons are important for the control of voluntary movements and behavioral processes including mood, reward, addiction, and stress (Chinta and Andersen 2005). The formation of intracellular inclusion bodies in the remaining neurons is a signature histopathological hallmark of PD. These inclusion bodies are spherical structures ranging between 8-30μm in size and are composed of amyloid filaments of the protein α-synuclein (Spillantini, Schmidt et al. 1997). These bodies were first identified by Fritz Lewy in 1912, and now bear his name as Lewy bodies (Goedert, Spillantini et al. 2013).

Although it has not became routine, detecting of functional and structural changes associated with PD progression is getting more feasible through the use of positron emission tomography (PET) imaging technology (Figure 1D). Various radiotracers were developed for various targets in the dopaminergic pathways including ^{18}F-fluoro-L-dopa (^{18}F-FDOPA), ^{11}C-dihydrotetrabenazine (^{11}C-DTBZ) which labels the vesicular monoamine transporter type 2 (VMAT2), ^{11}C-methylphenidate (^{11}C-MP) which labels the dopamine transporter (DAT), and ^{11}C-raclopride (^{11}C-RAC) which labels the dopamine D2 receptor. PET imaging opens the possibility to detect the pre-motor symptoms in the early stages and follow the progression of PD. (Stoessl, Martin et al. 2011)

The current treatments of PD are focused on improving of the life quality and functional capacity of the patients. L-DOPA (L-3,4-dihydroxyphenyalkanine) in combination with peripheral DOPA decarboxylase inhibitors is currently the most effective therapy for PD. The treatment can markedly suppress the symptoms of PD in the early stages. However, it loses efficacy after 5 to 10 years of treatment, while a number of complications, including motor fluctuation, dyskinesia, psychiatric and cognitive

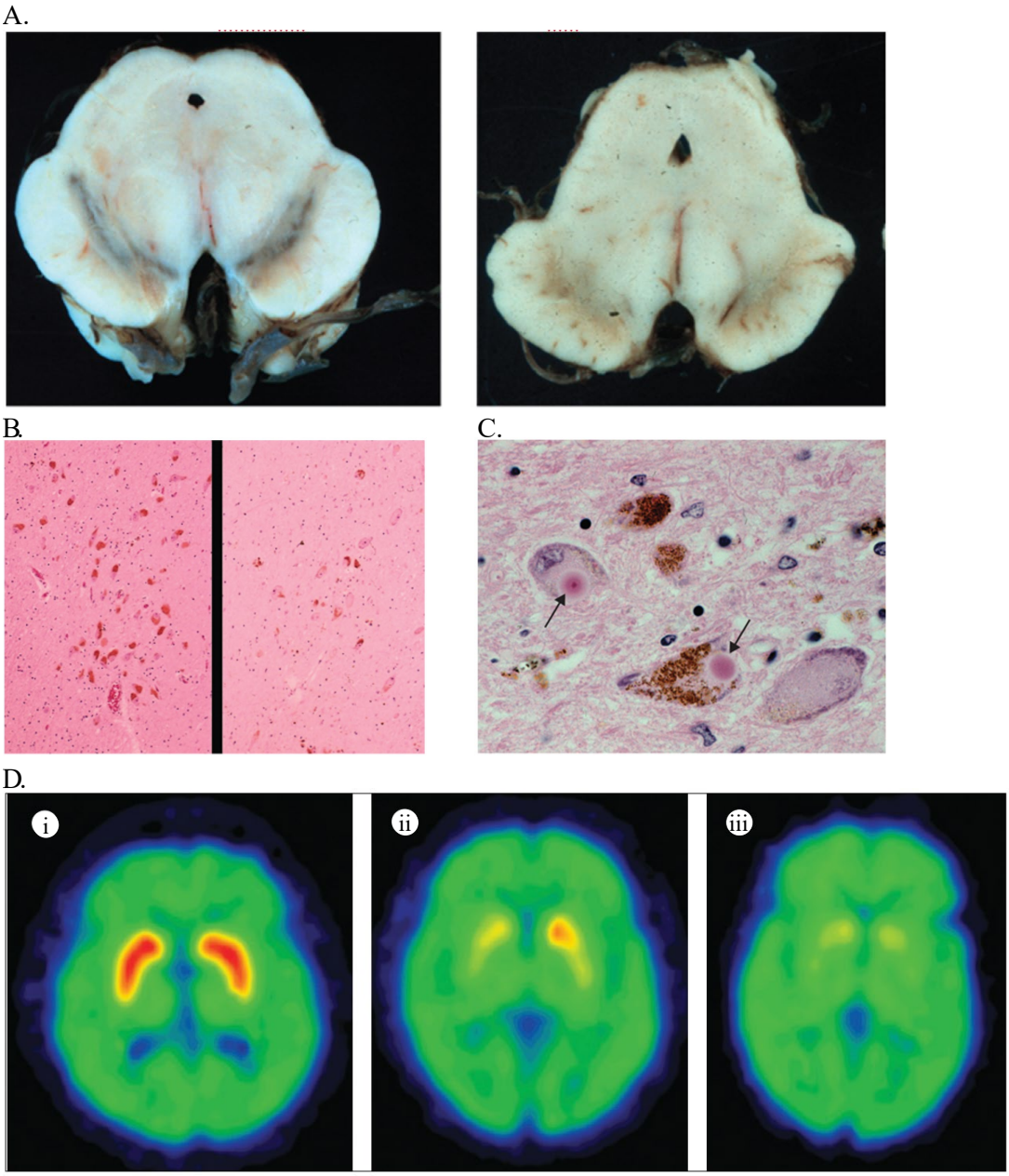


Figure 1. Neuronal degeneration in Parkinson's disease
A). Comparing basal ganglia from PD patient and normal control. The neuromelanin pigment of substantia nigra pars compacta (SNc) is reduced in the pathologic sample (right) versus the control (left). **B).** Reduced number of cells in SNc from the pathologic sample (right) versus the normal control (left). **C).** Lewy bodies in pathologic SNc specimen. **D).** Positron emission tomography (PET) imaging of dopamine transporter (DAT) labeled by ^{11}C-methylphenidate (^{11}C-MP) in: (i) a healthy individual, (ii) a clinically unaffected LRRK2 mutation (G2019S) carrier, (iii) a LRRK2 related PD patient. Pictures from review article "Advances in imaging in Parkinson's disease" on Lancet Neurology and "Harrison's principles of internal medicine" 19th edition (Stoessl, Martin et al. 2011, Kasper 2015)

disorders, will start to develop in the later stages of the disease (Varanese, Birnbaum et al. 2011). Without a mechanism-based therapy, there is still no cure and no means to effectively stop the progression of PD.

Epidemiology studies have shown that coffee drinking and cigarette smoking are associated with lower risk of PD, thus suggesting a protective effect (Hernan, Takkouche et al. 2002, Elbaz and Tranchant 2007). On the other hand, MPTP (1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine), a prodrug of neurotoxin MPP+ (1-Methyl-4-phenylpyridinium ion), was reported to selectively damage neuronal cells in the substantia nigra, and cause permanent PD-like symptoms (Langston, Irwin et al. 1984). A number of pesticides, including rotenone, paraquat, and organochlorine, were also reported to cause PD like symptoms in experimental animals; Thus, it suggests that exposure to these chemicals is likely to increase the risk of developing PD. (Fleming, Mann et al. 1994, Nandipati and Litvan 2016).

These studies shed light on the potential environmental factors that likely affect PD risk. Adopting a lifestyle with routine coffee drinking and avoiding exposure to PD-causative chemicals are likely helpful to prevent or reduce the chance of developing PD. However, environmental preventive strategies are not enough to cure/prevent PD, instead, developing a mechanism-based therapy is likely essential. Thus, understanding of the mechanism of the neuronal degeneration during the progression of PD is critical for therapeutic development.

In the last two decades, 18 genetic loci (listed in Table 1) were identified using linkage analyses, functional candidate gene approaches, genome-wide association studies, and exome sequencing in familial PD. These PD-associated genes are thus termed as the

PARK genes ranging from PARK1 to PARK18. (Sveinbjornsdottir, Hicks et al. 2000, Rocca, McDonnell et al. 2004, Klein and Westenberger 2012). Proteins encoded by these genes include α-synuclein (αSYN), Parkin, UCHL1, PINK1, DJ-1, LRRK2, ATP13A2, HTRA2, PLA2G6, FBX07, VPS35, and EIF4G1. The functions and cellular pathways of these proteins are likely linked to the mechanism of PD.

α-synuclein (α-Syn), encoded by the gene SNCA (PARK1, PARK4), is the main component of Lewy bodies. Missense mutation (PARK1) and duplication (PARK4) of the gene are found in rare familiar PD forms. Single nucleotide polymorphism within the promotor of the gene was found associated with sporadic PD. As mentioned previously, the formation of Lewy bodies is a major histopathological hallmark of all forms of PD and potentially causative to the neurodegeneration in PD.

α-Syn is a highly soluble protein that involved in synaptic vesicle release in healthy brain (Wang, Perovic et al. 2011). However, it is prone to misfold and aggregate in the β-sheet rich amyloid fibrils found in Lewy bodies (Cremades, Cohen et al. 2012). This form of protein aggregation is a shared feature in many neurodegenerative diseases including Alzheimer's disease and prion diseases. The misfolded protein fibrils act as a conformational template to promote the misfolding and aggregation of their cognate proteins into amyloid fibrils. (Spillantini, Schmidt et al. 1997, Goedert 2001) This effect was found intercellularly transmissive. After injection of α-Syn fibrils into the striatum of healthy mice, Lewy bodies were subsequently found in the brain and were distributed similarly as those found in PD. Dopaminergic cells in the substantia nigra of the same side of the injected striatum also showed progressive degeneration following the Lewy bodies formation, while the mice also showed reduction in motor function and co-ordination.

Thus, it suggests misfolded α-Syn fibrils have prion-like effect catalyzing the formation of Lewy bodies and that their presence is detrimental and leads to the degeneration of neurons (Luk and Lee 2014).

The misfolding and aggregation of α-Syn leads to a crisis of proteostasis (balance of proteins) in the cell, which is usually maintained by degradation and recycling though the protein sorting and vesicle transportation system with the trans Golgi and the endosome/lysosomal networks. Thus, dysfunction of this process is likely to interrupt the clearance of misfolded α-Syn, thus leading to the degeneration of neuronal cells.

Table 1. Parkinson's disease related genes

Gene Locus	Gene Product	Biochemical and Cellular function	Mode of inheritance	Status	Phenotype
PARK1 PARK4	α-synuclein, SNCA	Alpha synuclein is the main component of Lewy bodies. It is potentially neurotoxic and causative to cell death.	Autosomal dominant	Confirmed	Sporadic PD Early onset familial PD
PARK2	Parkin	E3 ubiquitin ligase. It is likely involved in mitochondrial stability and mitophagy. It is likely involved in the autophagy of depolarized mitochondria.	Autosomal recessive	Confirmed	Sporadic PD Early onset familial PD
PARK3	Unknown	-	Autosomal dominant	Unconfirmed	Late onset familial PD
PARK5	UCHL1	Ubiquitin C-terminal hydrolase, a de-ubiquitinating enzyme	Autosomal dominant	Unconfirmed	Late onset familial PD
PARK6	PINK1	PTEN-induced putative kinase, a serine/threonine kinase located in mitochondria. PINK1 activity results in recruitment of Parkin to mitochondria and induces mitophagy.	Autosomal recessive	Confirmed	Sporadic PD Early onset familial PD
PARK7	DJ-1	Remains not clear. Cysteine protease and oxidative stress sensor, but the protease activity seems lost.	Autosomal recessive	Confirmed	Early onset familial PD
PARK8	LRRK2	Cytosolic protein with multiple domains, has both GTPase and kinase activity. It is likely involved in the regulation of vesicles transportation in endosome, retromer and trans-Golgi network.	Autosomal dominant	Confirmed	Sporadic PD Late onset familial PD
PARK9	ATP13A2	A transmembrane endo-/lysosomal-associated P5 type transport ATPase. It is likely regulating endo-/lysosomal cargo sorting.	Autosomal recessive	Confirmed	Early onset atypical PD
PARK10	Unknown	-	Risk factor	Confirmed	Late onset familial PD
PARK11	Unknown	-	Autosomal dominant	Unconfirmed	Late onset familial PD
PARK12	Unknown	-	Risk factor	Confirmed	Late onset familial PD
PARK13	HTRA2	A serine peptidase was found in endoplasmic reticulum, mitochondria and nuclear. It is likely involved in apoptosis regulation.	Autosomal dominant or Risk factor	Unconfirmed	Late onset familial PD
PARK14	PLA2G6	An A2 phospholipase, it catalyzes the release of fatty acids from phospholipids.	Autosomal recessive	Confirmed	Early onset dystonia-parkinsonism

PARK15	FBX07	A F-box protein, it is a subunit of the ubiquitin protein ligase complex SCF (SKP1-cullin-F-box). It is likely on the downstream of PINK1 in mitophagy by targeting Parkin to the mitochondria.	Autosomal recessive	Confirmed	Early onset parkinsonian-pyramidal syndrome
PARK16	Unknown	-	Risk factor	Confirmed	Late onset familial PD
PARK17	VPS35	Vacuolar protein sorting (VPS) protein, it is a component of the retromer complex, which involves in retrograde transport of proteins from endosome to trans Golgi network.	Autosomal dominant	Confirmed	Late onset familial PD
PARK18	EIF4G1	Eukaryotic translation initiation factor 4 gamma 1. It is a large scaffolding subunit in the EIF4F complex, which elicit the recruitment of mRNA to ribosome.	Autosomal dominant	Unconfirmed	Late onset familial PD

B. LRRK2 related pathology of PD: dysfunction in cellular homeostasis results in aggregation of misfolded protein

Although the Lewy pathology theory provides an explanation on the mechanism of the neuronal degeneration in PD, it remains unclear how cells deal with the stress from the misfolded α-Syn under physiological conditions, and how this dysfunction occurs during the development of the disease. Since the discovery of α-Syn, there have been multiple PD-related genes identified that are functionally involved in autophagy, lysosome and vesicle transportation pathways (summarized in Table 1). These pathways, which mediates recycle of proteins through retromer-mediated retrograde vesicle trafficking and the degradation of proteins by endocytosis and autophagy, are essential for regulating the cellular homeostasis (Glick, Barth et al. 2010, Jovic, Sharma et al. 2010, Seaman 2012, Cuervo and Wong 2014). Thus, the PD-associated gene mutations may cause a dysfunction of the pathways that leads to accumulation of misfolded α-Syn and degeneration of neurons.

Mutations in the gene encoding leucine rich repeat kinase 2 (LRRK2) are common cause for PD. The gene locus of LRRK2, PARK8, was linked to inherited autosomal dominant form of PD in 2002 from a set of Japanese families (Funayama, Hasegawa et al. 2002). The underlying genetic cause was later identified in 2004 that mutations in the gene causes autosomal dominantly inherited PD with pathological features resembling sporadic forms of the disease (Paisan-Ruiz, Jain et al. 2004, Zimprich, Biskup et al. 2004). Addition to the familial PD cases, GWAS approaches also have identified LRRK2 as a risk locus that polymorphism in the gene region is associated with sporadic PD. (Satake, Nakabayashi et al. 2009, Simon-Sanchez, Schulte et al. 2009, Singleton and Hardy 2011). The common

effects of mutations in LRRK2 suggest a common pathogenic mechanism accounting for both inherited and sporadic PD.

The frequencies of the LRRK2 mutations are significantly high in some ethnic and regional populations, especially the G2019S mutation. In the Ashkenazi Jewish population, about 40 % of familial and 13 % of sporadic cases carry the G2019S mutation (Ozelius, Senthil et al. 2006). The G2019S mutation is also found in the north African Berber Arabs, where 39 % of familial and 40 % of sporadic patients carry the mutation (Lesage, Durr et al. 2006). In the Basque population, the frequency of the R1441G mutation in all patients is 22 %, which includes 46 % of familial PD and 2.5 % of sporadic PD cases. In contrast, the G2019S mutation has a low frequency in the Basque population, which includes only 2 % of patients, 1 % of familial and 2.5 % of sporadic PD cases (Gorostidi, Ruiz-Martinez et al. 2009). The high occurrence of LRRK2 mutations in PD patients suggests that the function of LRRK2 is essential for the disease. Autosomal dominant segregation pattern was found in these LRRK2-related PD cases, thus suggesting the pathogenic mutations are likely causing a "gain-of-function".

LRRK2 is likely involved in autophagy and lysosomal function (Roosen and Cookson 2016). LRRK2 is widely expressed in many different cell types, especially cells from kidney, lung and brain (Zimprich, Biskup et al. 2004, Maekawa, Kubo et al. 2010), thus suggesting its cellular function is not unique for neuronal cells. Subcellular localization studies showed that LRRK2 locate with endosomes, lysosomes and multivesicular bodies (MVBs) in cells and rodent brain (Biskup, Moore et al. 2006, Alegre-Abarrategui, Christian et al. 2009). Higher levels of lipdated autophagosomal marker LC3-II and lysosomal markers were found in LRRK2 KO cells and rodent kidneys (Tong,

Giaime et al. 2012, Baptista, Dave et al. 2013, Ness, Ren et al. 2013), thus suggesting the physiological function of LRRK2 is to regulate autophagosome formation and lysosomal maturation. Consistently, expression of pathogenic R1441C LRRK2 showed increased autophagic vacuoles in fibroblast cells (Alegre-Abarrategui, Christian et al. 2009), while the G2019S mutation showed both increased autophagic vacuoles and an accumulation of aggregated α-synuclein (Sanchez-Danes, Richaud-Patin et al. 2012, Reinhardt, Schmid et al. 2013). *In vivo*, transgenic mice carrying G2019S or R1441C mutation showed an accumulation of autophagic vacuoles in the cerebral cortex (Ramonet, Daher et al. 2011). Collectively, these data suggest that the pathogenic mutations in LRRK2 alter vesicular trafficking dynamics and cause a dysfunction of the autophagy-lysosomal pathway to degrade α-synuclein.

Consistently, LRRK2 was reported to interact with a number of proteins associated with vesicular transportation, including Rab7L1 (PARK16), Hsc70, and Rab10 (Dachsel, Taylor et al. 2007, MacLeod, Rhinn et al. 2013, Beilina, Rudenko et al. 2014, Steger, Tonelli et al. 2016). Rab7L1 (PARK16) is a Ras-related small G protein involved in the intraneuronal protein sorting in trans Golgi and endolysosomal networks. Deficiency of Rab7L1 was reported to cause neuronal degeneration in Rats (Helip-Wooley and Thoene 2004). Hsc70 (heat-shock cognate) is molecular chaperone that binds to clathrin and mediates the uncoating process from clathrin coated vesicles (CCVs). (Eisenberg and Greene 2007, Xing, Bocking et al. 2010). GAK (Cyclin G associated kinase) is a kinase that phosphorylates clathrin adaptors, while also serving as a cochaperone to clathrin, thus regulating the clathrin-mediated vesicle transportation (Zhang, Engqvist-Goldstein et al. 2005, Eisenberg and Greene 2007). Rab10 is a Ras-related small G protein involved in

asymmetric transportation of vesicles from the trans Golgi network to the plasma membrane in polarized cells. In neurons, it is involved in axonogenesis by regulating membrane vesicles trafficking to the axonal plasma membrane (Wang, Liu et al. 2011).

Collectively, these data suggest that LRRK2 is likely a regulator of vesicular transportation in protein sorting, endosomal, autophagosomal and lysosomal pathways. A dysfunction of these pathways would potentially disrupt cellular homoeostasis and cause degeneration of cells. However, it remains unclear how the pathogenic mutations disrupt the vesicular transportation pathways and lead to neurodegeneration in PD. Thus, understanding the function of LRRK2 and the biochemical effects of the pathogenic mutations is critical for the development of therapeutic strategies for PD.

C. ROC is a Ras-like but unique GTPase domain of LRRK2

The PD-associated LRRK2 is involved in signal pathways of vesicle transportation in endosome, autophagosome, lysosome and trans Golgi networks, thus regulating the homeostasis of proteins in cells. However, it remains unclear how LRRK2 biochemically regulating the pathways, and how the disease-associated mutations cause a "gain-of-function" of LRRK2 in the pathogenesis of PD.

The domain architecture of LRRK2. LRRK2 is a large complex protein with multiple domains and dual enzymatic activities, thus suggesting LRRK2 may be involved in multiple signaling pathways and not all domains are essential in the disease associated pathways. The LRRK2 contains at least 5 potential domains predicted based on the amino acid sequence, including: LRR (Leucine Rich Repeats), ROC (Ras-like GTPase domain of Complex protein), COR (C-terminal of ROC), Kinase domain, WD40 (tryptophan-aspartate repeats) (Cookson 2010). The N-terminal (of around 1 to 900 residues) prediction remains unclear, which is likely composed of ARM (ARMadillo repeats), ANK (ANKyrin repeats) or HEAT (Huntingtin, EF3, PP2A, TOR1 repeats) domains (Andrade and Bork 1995, Andrade, Petosa et al. 2001). Thus, LRRK2 is likely composed of total of 7 individual domains (Figure 2).

ARM, ANK, HEAT, LRR and WD40 repeats domains are common protein-protein interaction mediating domains. The ARM, ANK and HEAT repeats domains are composed of multiple α-helix rich tandem repeats which form large solenoid-like super-helix. Each repeat shares similar sequence and secondary structure arrangement, thus identifying them solely based on amino acid sequence is difficult (Andrade, Petosa et al. 2001, Chakrabarty and Parekh 2014). HEAT domain is commonly found in nuclear-cytoplasmic

transportation-associated proteins like importin-β and karyopherin-β (Chook and Blobel 1999, Tauchert, Hemonnot et al. 2016), while ARM domain is a common component in importin-α subunit (Goldfarb, Corbett et al. 2004). LRR is a domain composed of characteristic repetitive β-strand-loop-helix structure motifs rich in leucine residues, which also form curved solenoid structures that suitable for protein-protein interaction (Bella, Hindle et al. 2008). The WD40 domain is composed of tryptophan-aspartate-rich antiparallel β-strands which form a β-propeller structure with 7 blades, which are often involved in protein-protein interaction (Xu and Min 2011). Collectively, the ARM and ANK (or HEAT), LRR and WD40 domains are likely critical protein-protein interaction cores that mediates the substrate or regulator interactions of LRRK2.

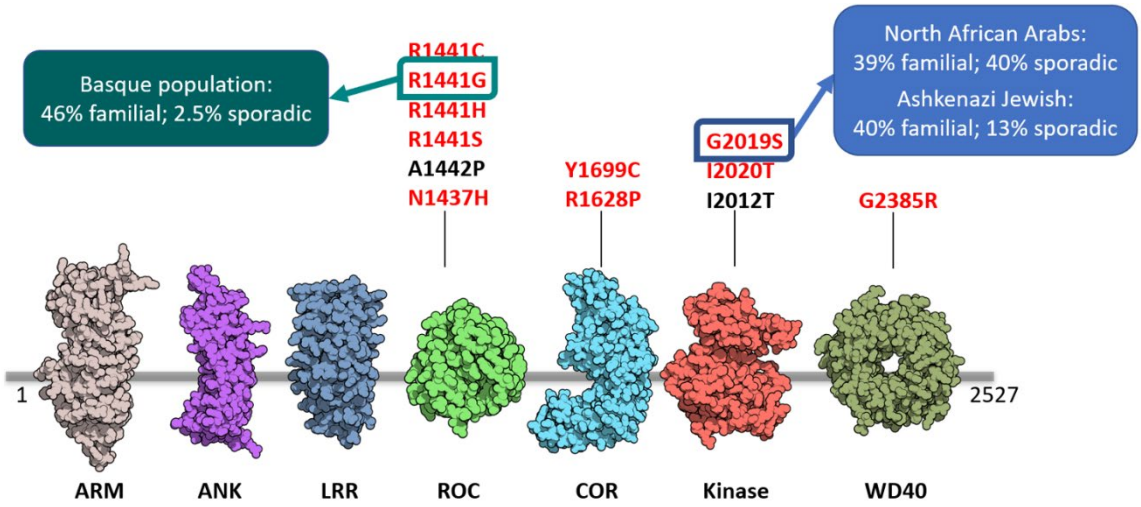


Figure 2. Leucine-Rich Repeat Kinase 2 (LRRK2): domains and pathogenic mutations
LRRK2 is a complex protein that consist of at least domains including leucine-rich repeat (LRR), Ras-like GTPase domain of complex protein (ROC), C-terminal domain of ROC (COR), serine/threonine kinase and WD40 domains, while the N terminal is likely contains 2 putative armadillo-repeats (ARM), and ankyrin-repeats (ANK) domains. The pathogenic mutations are mainly located in the GTPase (ROC) and kinase domains, thus suggesting their activity is likely critical in PD; While the mutations in COR and WD40 domains suggests they might regulate GTPase and kinase activity (Lesage, Durr et al. 2006, Ozelius, Senthil et al. 2006, Trinh, Farrer et al. 2006, Gorostidi, Ruiz-Martinez et al. 2009, Mata, Davis et al. 2017).

Surrounded by the protein-protein interaction domains, the ROC-COR-Kinase tandem domains form an enzymatic core region that holds both GTPase and kinase activities. So far, all prevalent pathogenic mutations that shown clear Mendelian segregation (autosomal dominant) are found in this region of the protein, which includes R1441C, R1441G, R1441H in the ROC domain, Y1699C in the COR domain and G2019S, I2020T in the Kinase domain (Zimprich, Biskup et al. 2004, Lesage, Durr et al. 2006, Trinh, Farrer et al. 2006, Gorostidi, Ruiz-Martinez et al. 2009). The clustering of pathogenic mutations in the ROC-COR-Kinase suggests this enzymatic core region is likely essential in the physiological and pathological function of LRRK2.

The next question is how pathogenic mutations alter the enzymatic activities of LRRK2. Early studies revealed that immunoprecipitated LRRK2 indeed has both kinase and GTPase activities *in vitro* (Guo, Gandhi et al. 2007). Following up studies revealed that pathogenic mutations (R1441C, R1441G) in ROC domain reduce GTPase activity (Guo, Gandhi et al. 2007, Lewis, Greggio et al. 2007, Li, Tan et al. 2007, Daniels, Vancraenenbroeck et al. 2011), while the mutation (G2019S) in the kinase domain results in increased kinase activity (Luzon-Toro, Rubio de la Torre et al. 2007). Thus, both GTPase and kinase activities are likely directly involved in the pathogenesis of PD. The up-regulated kinase activity, potentially as a threonine/serine MAPKKK-like kinase, is likely to cause an up-regulation of the downstream signaling of LRRK2, thus it is consistent with the "gain-of-function" hypothesis as suggested by the autosomal dominant segregation pattern. The reduced GTPase activity by mutations in the ROC domain, which is likely a Ras-like guanine nucleotide binding domain, is also consistent with "gain-of-

function" that a prolonged "active" state would cause increased activation to its downstream signaling interactors.

ROC is a Ras-like but unique GTPase domain. As a "Ras-like" GTPase domain, ROC is expected to have similar biochemical function and structural conformation. Ras super family is a protein family of small guanine nucleotide binding proteins. The founding members of the Ras super family were first discovered as oncogenes in the research of the murine sarcoma virus (MSV) by Jennifer Harvey in 1964 (Malumbres and Barbacid 2003). And later other small G protein families that has functional similarity to the Ras are classified together into the Ras super family, which now includes Ras, Rho, Rab, Ran and Arf families. Various Ras family small G proteins involve in different subcellular environments and interact with various regulators (GAP, GEF, GDI, etc) and effectors, so that they modulate a diverse range of cellular signaling. However, they share common biochemical characteristics and commonly act as binary molecular switches in the signaling pathways (Colicelli et al. 2004; Wennerberg et al. 2005). The biochemical activity of small G proteins is to hydrolyze GTP (guanosine triphosphate) to GDP (guanosine diphosphate). However, the important biological function of small GTPase is more than the biochemical catalysis itself. As binary switches in cellular pathways, Ras family proteins turns signals ON and OFF by the cycles of GDP and GTP binding states (Figure 3B). The switch turns ON when GTP is bound, which enables the interaction to downstream effectors. While the switch turns OFF after the GTP was hydrolyzed into GDP, and it stays OFF before the next binding event of GTP. During the switch OFF state, effectors dissociate and inactivate the pathway.

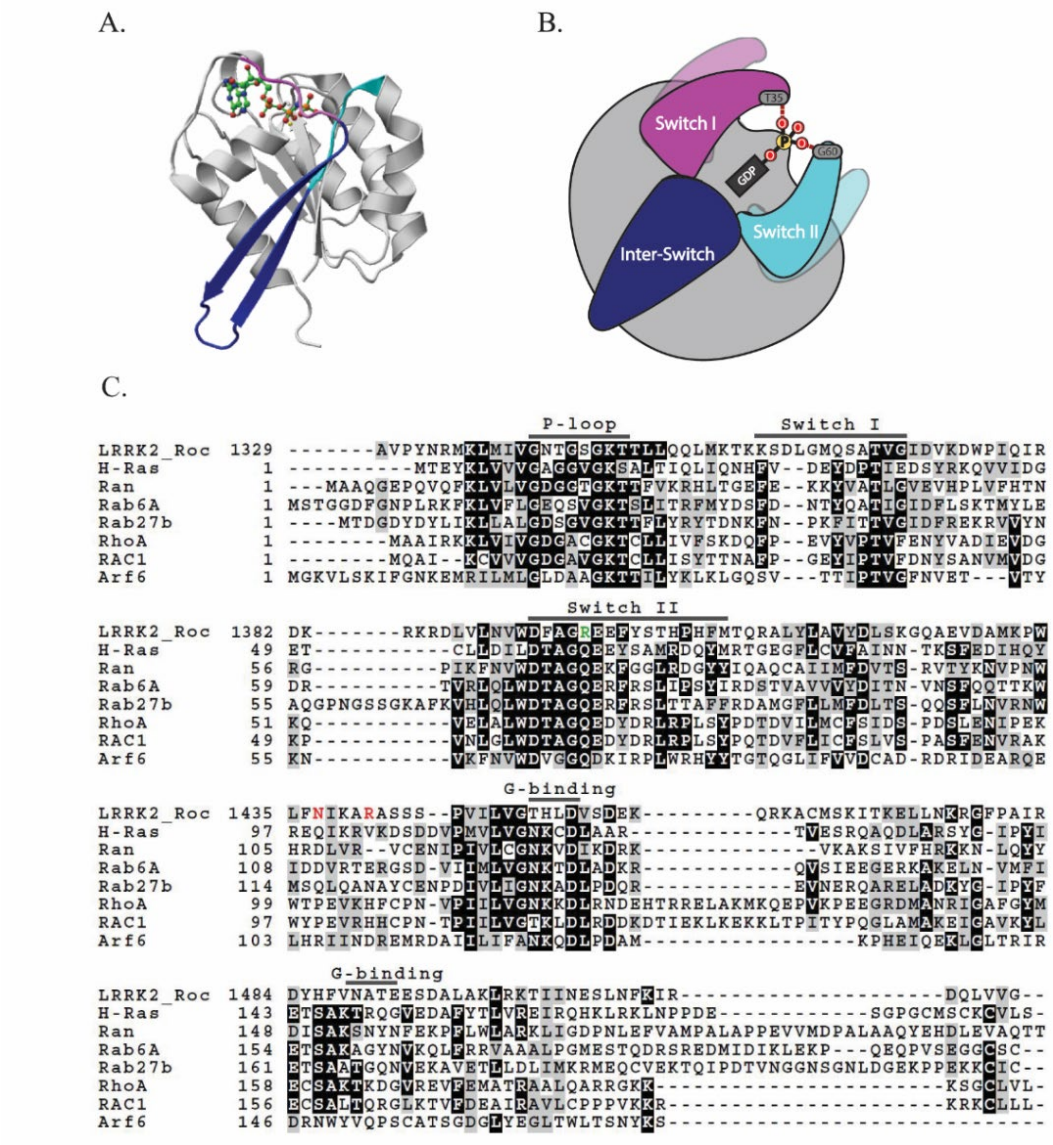


Figure 3. Structural and functional motifs of Ras family small G proteins
A). A typical structure model of a Ras family small GTPase (Rab6, 1YZQ) bound with GTP analog GppNHp. The switch I, switch II and inter-switch motifs are labeled as purple, cyan and blue; **B)** Diagram that describes the conformational change of the switch I, switch II motifs during the cycles of GDP and GTP turn-over. When GTP is bound, the switch I and switch II motifs form hydrogen bond with the oxygens of the γ phosphate of GTP, thus taking a "stretched" conformation. After the GTP is hydrolyzed to GDP, where the γ phosphate is absent, the switch motifs adopt a "relaxed" conformation. Thus, the small G protein cycles between the GTP-bound "active" state and GDP-bound "inactive" state as a molecular switch **C)** Alignment of ROC domain with typical G proteins in the Ras superfamily. ROC contains conserved functional motifs comparing to other G proteins. However, the pathogenic mutation sites (N1437, R1441 labeled in red) are in the varied regions without corresponding residue in other G proteins, thus suggesting the disease-associated mechanism is likely unique to ROC. Another unique feature of ROC is the R1398 residue (labeled green) in the switch II region, where a potential protective mutation (R1398H) was found. In other G proteins the residue is a highly conserved glutamate instead of arginine.

The molecular switch function of G proteins is mediated by the conformational plasticity of several structural motifs near the catalytic site of the enzyme. These conserved functional motifs were thus termed as the G motifs (Figure 3C): G1 (P-loop), G2 (switch I), G3 (switch II), G4 and G5 (G-binding). The G1 motif: GXXXXGKS/T is the phosphate-binding loop (P-loop) that interacts with the β, γ phosphate of the guanine nucleotide. The G4 motif: T/NKXD interacts with the nucleotide base of the guanine nucleotide. Thus, the G1 and G4 are the main contributors to the binding of guanine nucleotide. The G2 motif: T and the G3 motif: DXXGQ/H/T are the switch regions that form hydrogen bond with the oxygens of the γ phosphate of GTP, and thus taking a different conformation in the ON state (GTP bound form) and OFF state (GDP bound form) of the G protein. While in the G5 motif: C/SAK/L/T, the invariant alanine interacts with the guanine oxygen, thus provides specificity to guanine instead of adenine (Vetter and Wittinghofer 2001). The conformational change of two loop regions, switch I and switch II, plays a critical role in the biological function of small G proteins as molecular switches in the cellular pathways. When the catalytic site is bound with GTP, the switch regions interact with the gamma phosphate of the GTP and forms a conformation close to the catalytic site (Figure 3A and 3B). This conformation of the switch regions mediates the interaction of the G protein with the downstream protein interactor, thus turning ON the molecular switch. While GDP is bound in the catalytic site, the switch regions cannot interact with GDP and instead form a conformation distant from the catalytic site (Figure 3B). In this conformation, the G protein cannot interact with the downstream interactors. Thus, the signaling switch is OFF. The conformational change of the switch regions is critical for the molecular switch function of GTPase.

However, ROC contains unique disease-related residues that are not present in Ras and other small G proteins. The pathogenic mutations in the GTPase domain are especially focused around residue R1441, including R1441C, R1441G, R1441H, R1441S and nearby potential risk factor N1437H and A1442P (Zimprich, Biskup et al. 2004, Zabetian, Samii et al. 2005, Trinh, Farrer et al. 2006, Gorostidi, Ruiz-Martinez et al. 2009, Mata, Davis et al. 2016). Comparing with other major classes of G proteins in the Ras superfamily (Figure 3C), these residues are located in non-conserved regions, and there is no residue corresponding to the PD-related residue R1441. Thus, it suggests the R1441-associated pathogenic mechanism is likely unique to ROC. Interestingly, the R1398 residue site of a PD-protective mutation R1398H in the switch II motif is different from the conserved glutamine residue as in other G proteins. In other G proteins, the glutamine residue helps stabilize a nucleophilic water near the catalytic site and thus act as an enhancer for the GTPase activity. Thus, the R1398 may also mediate a nucleophilic water at the catalytic site, and the protective mutation may alter the GTPase function by affecting this interaction. Collectively, albeit the conserved functional motifs, ROC contains unique features that may mediate functional mechanism different from other Ras-like small G proteins, but likely similar to other complex protein homologs that contains ROC domain.

ROCO protein family is a conserved complex protein family which contains the ROC-COR tandem bidomain. Indeed, Phylogenetic analysis of different ROC domains revealed that the ROC-COR tandem bidomain arrangement is conserved in a monophyletic group of complex proteins, which was termed as the ROCO protein family (Bosgraaf and Van Haastert 2003). A number of proteins were classified in the ROCO protein family, including DAPK, LRRK1, GbpC, MASL1, QkgA and Pats1 (Lewis 2009,

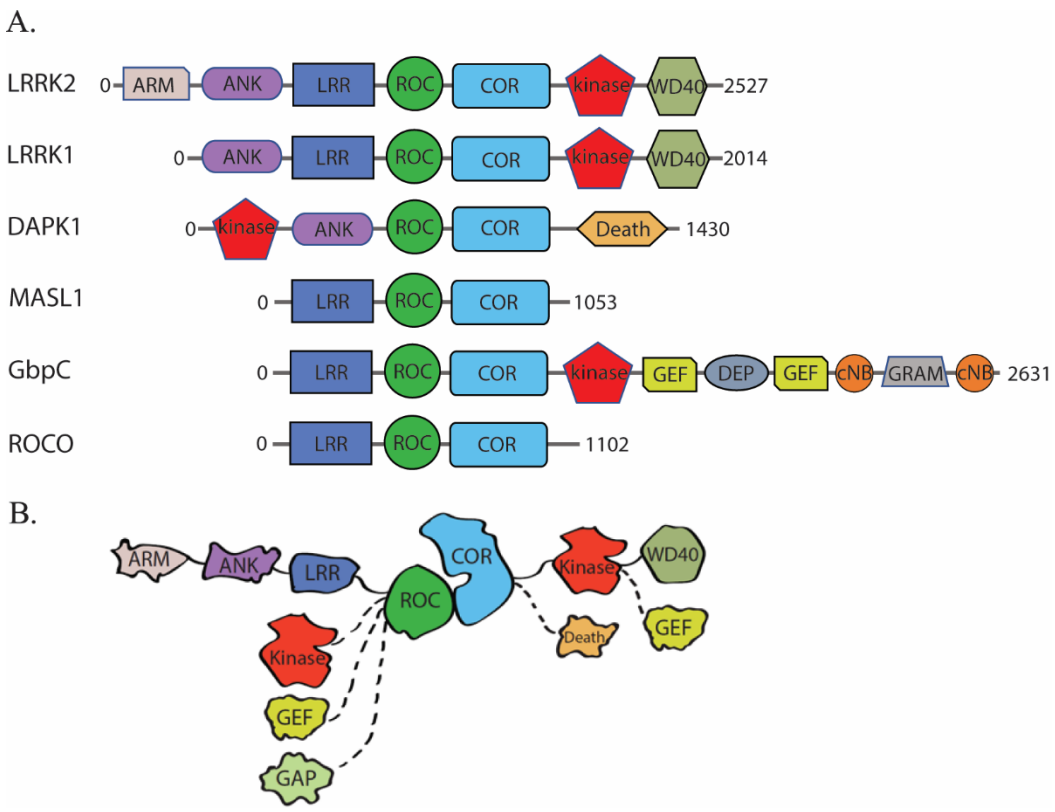


Figure 4. ROC-COR bidomain is a conserved functional unit in ROCO protein family
A). Diagram showing the recombined arrangement of ROCO family proteins, including human LRRK2, LRRK1, DAPK1, and GbpC from *D. discoideum*, ROCO from *C. Tepidum*. The ROC-COR tandem domains is conserved among different ROCO family proteins, while other domains vary, including ARM, ANK, LRR, kinase, WD40, death, cNB (cyclic nucleotide binding), DEP (disheveled), GEF (guanine nucleotide exchange facor), GAP (GTPase activating protein), GRAM (glucosyltransferase) domains (Lewis, P.A. Biol Cell. 2009). **B).** Comparing the domains arrangement in LRRK2 (connected with solid lines) to other alternative domain arrangements (connected with dashed lines) found in ROCO protein family. The conserved ROC-COR tandem domain arrangement suggests they are likely one functional unit, while the other varied domains are likely interacting with ROC-COR through different mechanisms. Figure adapted based on Civiero L, Chem Biol 2014.

van Egmond and van Haastert 2010). By far, the ROCO proteins were found widely in various species including thermophilic green sulfur bacteria (*Chlorobaculum tepidum*), archaea (*Methanosarcina barkeri*), placozoa (*Trichoplax adhaerens*), amoebozoan slim mold (*Dictyostelium discoideum*), thale cress (*Arabidopsis thaliana*), roundworm (*Caenorhabditis elegans*), fruit fly (*Drosophila melanogaster*), Zebrafish (*Danio rerio*), and human beings (Civiero, Dihanich et al. 2014).

The ROC-COR tandem bidomain is recombined with various other domains in the ROCO protein family (Figure 4A), including leucine rich repeats (LRR), ankyrin repeats (ANK), kinase, WD40 repeats, death, cNB (cyclic nucleotide binding), DEP (disheveled), GEF (guanine nucleotide exchange factor), GAP (GTPase activating protein), GRAM (glucosyltransferase) domains. These other domains are variably arranged either on the N- or C-terminal of the conserved ROC-COR tandem domains (Figure 4B) (Marin, van Egmond et al. 2008, Lewis 2009, van Egmond and van Haastert 2010, Civiero, Dihanich et al. 2014), thus suggesting the ROC-COR tandem bidomain is likely a functional unit that regulates the various other domains. Interestingly, the GTPase activity of *C. Tepidum* Roco protein (ctROCO) was reported depending on the dimeric conformation mediated by the COR domain. Mutations in the dimeric interface disrupted the dimeric conformation and cause reduced GTPase activity, thus suggesting a regulatory role of the COR domain mediated dimerization on GTPase function (Gotthardt, Weyand et al. 2008, Sen, Webber et al. 2009). Consistently, the Y1699C mutation in the COR domain of LRRK2 was found pathogenic and causing reduced GTPase activity, thus suggesting a potential regulatory role of COR domain on the GTPase activity of the ROC domain (Daniels, Vancraenenbroeck et al. 2011). Thus, we decided to investigate on both the ROC domain and COR domains of LRRK2 to analyze its detailed biochemical and biophysical properties.

D. Rationale and overview of this chapter

Pathogenic mutations in the GTPase domain of LRRK2 (ROC, Ras-like GTPase domain of Complex proteins) are common genetic causes of PD. The R1441 residue is a hotspot of pathogenic mutations, including R1441C, R1441G, R1441C, R1441H (and recently reported R1441S), thus suggesting the residue is critical to the function of LRRK2. However, it remains unclear how the pathogenic mutations at residue R1441 alter the biochemical function of the ROC domain.

The overall goal of this chapter is to examine the biochemical and biophysical properties of the ROC-COR bidomain, and determine the functional alteration caused by the prevalent pathogenic mutants in ROC (R1441C, R1441G and R1441H). To address this, we have made a number of recombinant protein constructs to stably express ROC-COR, however only a few constructs are soluble and stable enough for purification. The major two stable constructs used in this study is ROC (1329-1520) and ROC-cor1 (1329-1580).

The stable ROC (1329-1520) construct has a complete residue coverage of an isolated ROC domain, which shows both dimeric and monomeric forms in solution. Different from Ras family small GTPases which are functional monomers, we have observed a dynamic dimer-monomer equilibrium of ROC in solution, which is driven by GDP and GTP binding. During the interconversion, dimeric ROC turns into monomeric form after GTP binding, while it partially converts back to dimeric conformation after incubation of GDP. Thus, it suggests the dynamic dimerization equilibrium is likely related to it GTPase activity. ROC is an active but slow GTPase in solution, which has a low intrinsic activity similar to other G proteins. However, ROC has a lower binding affinity

to guanine nucleotides comparing to Ras, thus suggesting that ROC might have a different nucleotide exchange mechanism independent to guanine nucleotide exchange factors (GEFs). Given the unique GDP/GTP binding coupled dimer-monomer interconversion, the nucleotide exchange mechanism is likely directly associated with conformational changes of the dimeric interactions.

The extended construct ROC-cor1 (1329-1580) contains a N-terminal fraction of COR (cor1 1521-1580), which is likely directly interact with ROC. Thus, it was expected to shed light on the function of ROC-COR bidomain. Different from construct ROC, ROC-cor1 was only observed as monomeric form in solution, while GDP treatment cannot convert it to dimeric form, thus suggesting the extended cor1 may interact with the monomeric ROC domain. Consistently, binding assay reveals it has enhanced binding to GTP than construct ROC, thus suggesting a stabilized GTP-bound monomeric form of the ROC domain. The kinetics of the GTPase activity showed lower Michaelis constant K_M than construct ROC, thus consistently suggesting an enhanced GTP binding as the substrate. However, ROC-cor1 has a remarkably reduced K_{cat} and thermal stability (a 4°C decrease of melting temperature). Thus, ROC-cor1, which contains a partial fragment of COR domain, might cause a highly unstable dimeric conformation in solution, thus causing a disruption to the GTPase activity.

To investigate the biochemical and biophysical consequences of the pathogenic mutations, we used the most stable construct ROC (1329-1520) as the template for mutagenesis. Comparing the wild type ROC to the pathogenic mutants R1441C, R1441G and R1441H, the GTPase activity is dramatically decreased. Detailed comparison of the Michaelis-Menten kinetics between the wild type and mutated ROC showed consensually

reduced K_{cat}, thus suggesting all these pathogenic mutations reduce the intrinsic hydrolysis of GTP. However, the Michaelis constant K_M is either not dramatically changed (R1441G), decreased (R1441H) or increased (R1441C), thus suggesting that mutations have different effects on substrate accessibility. Further binding assay showed a trend consistent to the Michaelis constant K_M, in which R1441H has tighter binding, R1441C showed weakened binding, and R1441G is similar to the wild type. All these pathogenic mutations at R1441 were found to disrupt the formation of the dimeric conformation in solution. Further dimer-monomer conversion assay also showed the pathogenic mutant R1441G monomer was not converted to dimeric form after GDP incubation. Thus, it suggests the coupled GTPase cycles and dynamic dimer-monomer interconversion are disrupted by the R1441 pathogenic mutations. Thus, the biochemical consequence of the pathogenic mutations is causing a prolonged GTP-bound monomeric form of ROC, thus leading to enhanced signals to the downstream effectors and a "gain of function" to the disease-associated signaling pathway.

1.2 Materials and Methods

A. Multiple species sequence alignment and secondary structure prediction

The multiple species alignment was done by Clustal2.1 (http://www.clustal.org). To design potential stable construct, the amino acid sequences of the human LRRK2 was aligned with those from rat, mice, swine, bovine and zebra fish. To compare the functional and structural motifs, the amino acid sequencing of ROC is also aligned with other small G proteins, including H-Ras, Ran, Rab6A, Rab27b, RhoA, RAC1 and Arf6. The secondary structure of the LRRK2 was predicted by the protein secondary structure prediction server, Jpred4 (http://www.compbio.dundee.ac.uk/jpred/).

B. Homology Modeling

Homology structure models of the ROC domain monomer and the ROC-COR tandem domains were generated using the program Modeller 9.19 (Andrej Sali, UCSF, San Francisco, CA). The ROC-COR bi-domain model was built based on the *C. tepidium* ROCO protein (3DPU) structure (Gotthardt, Weyand et al. 2008). Alternatively, some model is also prepared using SWISS-MODEL (https://swissmodel.expasy.org/), which was used to analyze in a wider homology-modelling candidate template pool (Waterhouse, A. et al. 2018; Bienert, S. 2017; Guex, N. 2009; Benkert, P. 2011; Bertoni, M. 2017).

C. Limited protease digestion followed by mass spectrometry

To help decide the terminals for the design of the recombinant GTPase domain constructs of LRRK2, the full length LRRK2 protein was digested with limited time and amount of trypsin. The strategy is that the relative rigid globular folded domains are more resistant to protease digestion than the flexible “disordered” loops, so that the regions boundaries could be identified by the followed mass spectrometry.

The LRRK2 full length protein was prepared as 1 mg/ml and digested with 10 μg/ml trypsin in buffer (total volume of 100μl) containing 10 mM Tris-HCl, pH 7.8, 0.1 mM NaN_3. The digestion mix was then incubated in 37˚C to start the protease reaction. 10 μl sample was taken out of the reaction mix each time point at 0, 1, 3, 5, 8, 10, 20, 30 min. The samples were immediately mixed with the SDS PAGE loading dye and boiled for 3 min to stop the protease reaction, then stored in -20 ˚C. After all samples were ready, the samples were run on the SDS PAGE. To identify the protein identity in each of the bands, the gel sample was cut out at each of the clear bands on the SDS PAGE gel and set out to mass spectrometry analysis.

D. Cloning the recombinant ROC domain and expression tests

The recombinant ROC domain DNA was PCR amplified from the existing pFL-LRRK2 plasmid previously prepared in the lab. The amplified DNA fragment of each designed recombinant ROC domain was then cloned into either pET-Duet-1 (N-terminal 6 × histidine tag, Novagen), pGEX6p1 (N-terminal GST tag, GE health care) or pET28b-SUMO (N-terminal 6×histidine tag followed by sumo, generated based on pET28b) vectors for protein expression. The sequence of the plasmids was verified by DNA sequencing (Genscript). The plasmids were transformed into Rosetta2 (DE3) cells for protein expression.

To test the expression level in small scale, 5 ml of the cell culture was prepared and culture in 37˚C to OD600 of 0.8. The protein expression was then induced with 0.5 mM IPTG. The cells were then continue cultured in 20˚C overnight for protein expression. To lyse the cells, the harvest cells pellet was resuspended in 200 μl of the B-PER bacterial protein extraction reagent (Thermo Fisher) and then sonicated for 15 sec. To analyze the

solubility of the expressed protein, the cells lysate was centrifuged at 16000 × g, 10 min to separate the soluble and insoluble fractions. The soluble and insoluble protein samples were then run on SDS PAGE.

The recombinant protein expressing cells were resuspended in buffer containing 30 mM HEPES, 250 mM NaCl, 10 mM $MgCl_2$, 10 mM Glycine, 10 mM imidazole, 10 % glycerol. To lyse the cells, the resuspended cells solution was sonicated 5 min with pulses set as 3 sec on, 12 sec off. The lysate was then ultra-centrifuged at 35000 rpm, 1 hour to collect the soluble fraction. The soluble lysate was then incubated with Ni-NTA affinity beads (QIAGEN) in 4˚C 2 hours. The Ni-NTA beads were washed 3 times with 5 bed volume of the buffer containing 30 mM HEPES, 250 mM NaCl, 10 mM $MgCl_2$, 10 mM Glycine, 40 mM imidazole, 10 % glycerol. The protein was eluted from the Ni-NTA beads with the buffer 30 mM HEPES, 250 mM NaCl, 10 mM $MgCl_2$, 10 mM Glycine, 300 mM imidazole, 10% glycerol. To further polish the purity, the protein solution was then concentrated and injected to the size exclusion column (Superdex200 16/60, GE) on the AKTA FPLC system (GE Healthcare) with flow rate 0.5 ml/min, 2 ml each fraction, end volume 140 ml of final buffer containing 30 mM HEPES (pH 7.4), 0.15 M NaCl, 10 mM $MgCl_2$, 10 mM Glycine, 1 mM DTT, and 10 % Glycerol. The fractions containing the target protein was combined and concentrated to around 20 mg/ml. For long term storage, the protein solution was frozen in liquid nitrogen and stored in -80˚C.

E. Size exclusion chromatography coupled with multiangle light scattering (SEC-MALS)

To determine the oligomeric states of ROC in solution, Multiangle Light Scattering (MALS) was used to test the absolute molecular weight of the samples in soltuion. The

MALS was setup tandem with a silicon-based size exclusion chromatography (SEC) column (WTC-030S5, Wyatt Tech) on the AKTA FPLC system (GE healthcare). The differential refractometer Optilab T-rEX (Wyatt Tech) and multiangle light scattering detector Dawn Heleos II (Wyatt Tech) are tandemly connected on the downstream of the FPLC system. Thus, the injected sample first got resolved into single peaks based on the size in SEC; the concentration was then determined based on UV absorbance detected by AKTA FPLC and differential refractive index (dRI) measured by Optilab T-rEX; the light scattering signal on 18 different angles was then determined by the Dawn Heleos II. The samples for each measurement contained about 1 mg of purified ROC sample in buffer of 30 mM HEPES (pH 7.4), 0.15 M NaCl, 10 mM $MgCl_2$, 10 mM Glycine, 1 mM DTT, and 10 % Glycerol. The flow rate of the FPLC was set to 0.4 ml/min, and data acquisition interval of MALS was set to 1 second. The data analysis and molar mass calculation were processed in the ASTRA software (Wyatt Tech). The graphs showing the SEC elution profile and the calculated molar mass of each peak were plotted in Graphpad Prism (Graphpad software).

F. Dimer-monomer inter-conversion assay

To determine how the ROC converts between dimeric and monomeric forms in response to GTP/GDP incubation, the ROC samples of 15 mg/ml (equal to 0.64 mM) were prepared by incubating with 16 mM (25× concentration of ROC) different guanine nucleotides in buffer 30 mM HEPES (pH 7.4), 0.15 M NaCl, 10 mM $MgCl_2$, 10 mM Glycine without addition of EDTA. To test the dimer to monomer conversion, the sample was incubated in room temperature with 16 mM of either GDP or GTP for 6 hours. To test the monomer to dimer conversion, the monomeric protein samples were first obtained by

GTP incubation and desalting purification (Zeba spin desalting column, Thermo Fisher). The monomeric samples were then incubated with 16 mM GDP or GTP for 6 hours. The ratios of monomeric and dimeric forms of the incubated samples were then determined by size exclusion chromatography (Superdex 75, GE Healthcare) on the AKTA FPLC system (GE Healthcare).

G. Circular dichroism spectroscopy

To measure the circular dichroism (CD) spectra, the protein sample was prepared as 1.0 mg/ml in buffer containing 30 mM Tris-HCl pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT and 5 % Glycerol. The protein solution was then loaded into a 0.2 mm cuvette. The data was then collected in room temperature on the MOS-450 AF/CD spectrometer (Biologic Science Instruments). The spectrum was scanned from 190 nm to 250 nm with data acquisition time of 1.0 s. To calculate the percentage of the secondary structure elements, the data sets were analyzed on the dichro-web server (http://dichroweb.cryst.bbk.ac.uk/html/userguide.shtml), where the algorithm "CONTIN" and reference data "SET7" was used in the analysis.

H. Fluorescence polarization nucleotide-binding assay

To analyze the binding affinity of the guanine nucleotides, the fluorescence polarization assay was used. The guanine nucleotides GDP or GppNHp (GTP analog) were labeled with fluorophore BODIPY (excitation at 485 nm and emission at 535 nm). When guanine nucleotides are in their free form, the emitted light after the excitation is depolarized due to the rapid rotation of the molecules. On the other hand, when the guanine nucleotides are bound to the ROC protein to form a big complex, the rotation is much

slower, and the emitted light will be more polarized. Thus, the fluorescence polarization can be used measure the binding of the guanine nucleotides to the ROC protein.

The fluorescence polarization signal was measure using the EnVision 2102 multiplate reader (Perkin Elmer). To measure the binding affinity, the BODIPY-GDP (150 nM) or BODIPY-GppNHp (100 nM) was titrated with the ROC protein from 0.1 μM to 15 μM where saturation was reached. The measurements were performed at 25˚C in buffer containing 30 mM HEPES (pH 7.4), 150 mM NaCl, 10 mM MgAc, 10 mM Glycine, 4 mM EDTA, 1 mM DTT, 10 % Glycerol. The data were analyzed using GraphPad Prism 6 (GraphPad Software).

I. GTPase activity assay

The GTPase activity was measured using the EnzChek phosphate detection assay kit (Thermo Fisher). The EnzChek phosphate detection assay is a spectrophotometric method for the quantification of inorganic phosphate that released from GTPase reactions. In the presence of inorganic phosphate, the 2-amino-6-mercapto-7-methylpurine riboside (MESG, maximum absorbance at 330 nm) substrate was converted by the purine nucleoside phosphorylase (PNP) to the ribose 1-phosphate and 2-amino-6-mercapto-7-methylpurine (maximum absorbance at 360 nm). Therefore, the GTPase reaction was accessed by measuring the absorbance at 360 nm. The ROC protein (30 μM) was incubated with 2 mM GTP in reaction buffer containing 30 mM HEPES, 150 mM NaCl, 10 mM $MgCl_2$, 10 mM Glycine, 1 mM DTT, and 10 % Glycerol at 25˚C. The absorbance at 360 nm was measured in every 3 min for a total duration of 3 hours on a SpectraMax microplate reader (Molecular Devices). The amount of the inorganic phosphate release at each time

point was calculated using the phosphate standard curve. The data analysis was done with the GraphPad Prism 6 (GraphPad Software).

J. Thermofluor assay

A fluorescence based thermal shift Thermofluor assay (Thermo Fisher) was used to test the melting temperature of the ROC domain protein. The fluorescent dye, Sypro Orange (Thermo Fisher) binds to hydrophobic patches of the protein. Therefore, when temperature is increased to the melting point that protein unfolds, an increased fluorescence signal can be observed due to the increased binding of Sypro Orange dye to the denatured protein. The 5000 × Sypro orange (Thermo Fisher) stock solution was diluted to 10 × working solution.

To measure the melting temperature, 12.5 µl of the 10× Sypro Orange was mixed with 12.5 µl of the ROC protein on a 96-well thin-wall PCR plate in buffer containing 30 mM HEPES pH 7.4, 0.15 M NaCl, 10 mM MgAc, 10 mM Glycine, 1 mM DTT, 4 mM GDP or GTP non-hydrolysable analogue GppNHp, 2 µM LMNG and 10 % Glycerol. The melting curve was measured in the Real Time PCR Detection system, Mastercylcer Realplex (Eppendorf) from 20 to 85˚C. The fluorescence signal at wavelength 550nm was measured each temperature increment of 0.4˚C.

1.3 Results and Discussion

A. Construction of a stable and active recombinant protein of LRRK2

To study the detailed mechanism of the GTPase function of LRRK2, we need to obtain stable constructs of the ROC or ROC-COR tandem domains to purify protein samples with sufficient amount and purity. Especially, for the determination of the detailed biochemical kinetics and the molecular structure, milligrams level amount and better than 90 % purity is often required. Thus, the first part of this chapter is to obtain stable and high yield recombinant constructs, and suitable procedures to purify the protein sample.

Strategies for designing LRRK2 recombinant protein constructs. LRRK2 is a large protein that consists of 2527 amino acids. It contains 7 individual domains, thus suggesting the determination of the domain boundaries is critical. Here, we adopted 4 major strategies to help find potential stable constructs: The first strategy is to utilize the secondary structure prediction on LRRK2 amino acid sequence (Figure 5C.). The secondary structures: alpha helix, beta strands are the building blocks for peptide chain folding and are often critical for the stability of the protein tertiary structure. Thus, it is likely making the constructs more stable in solution when the helixes and strands are intact. The construct boundaries P1261, E1301, D1307, A1329, G1520, N1540 were designed accordingly. The second strategy is to align the LRRK2 amino acid sequence to multiple homologues from other species. This approach is to identify the critical functional or structural motifs which are usually conserved throughout evolution. Thus, the construct boundaries were designed in the variable regions so that the functional motifs are not disrupted. Accordingly, the construct boundaries Y1277, E1301, N1504 were designed. The third strategy is the homology structure modeling. The structure of protein homologues

often has similar folding topology. At the time in 2011, the only available homologue structure of LRRK2 is the ROC-COR tandem domains structure (3DPU) of ROCO (25 % sequence identity) from *C. Tepidum* (Gotthardt, K. et al. 2008). Based on this structure model, the software Modeller (Fiser and Sali 2003) is used for the comparative structure modeling. Based on the homology model, the COR domain is folded in 3 relatively independently folded regions (Figure 5A). Thus, the constructs were also designed to contain ROC with a fragment of COR. The terminals Q1580, D1587, K1655, Q1657, K1832, Q1845, I1850 were designed accordingly. The last strategy is using limited protease digestion to identify the flexible loop regions in LRRK2, which is often in "disordered" conformation in solution. The extended flexible loops are prone to be digested by the low concentration protease treatment, while the more rigid globular domains are more resistant to the treatment and likely stay intact. The digested sample was then separated by SDS PAGE. To identify the digested fragments, the bands were cut out and tested using mass spectrometry. The resistant fragments indicated internal rigid folded regions, while the digestion sites were likely in the extended loops regions. Accordingly, the boundary residues K1256, K1499, K1512 were selected based on this method.

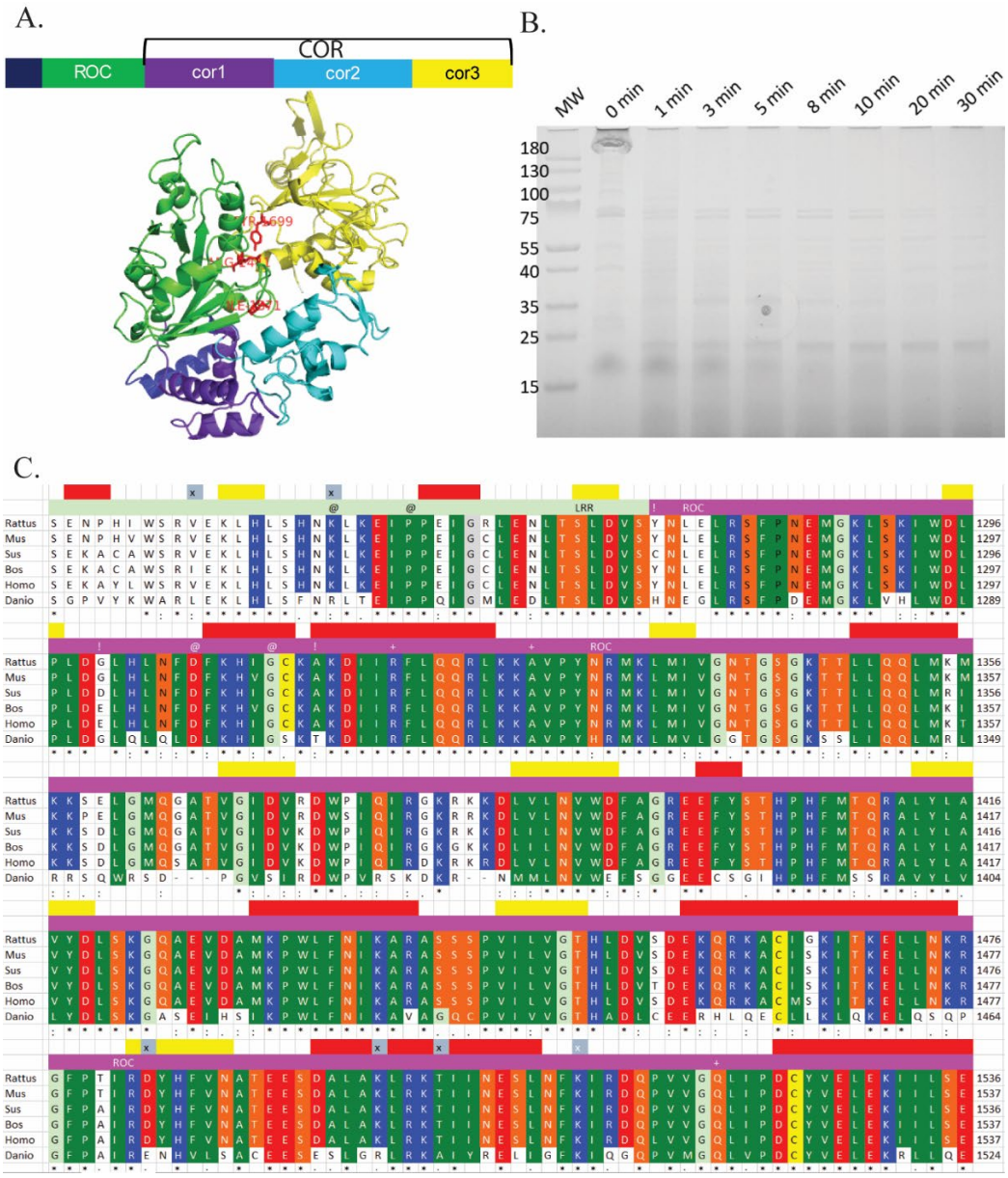


Figure 5. Strategies for construct design of the ROC domain of LRRK2
A). Homology model of the ROC domain based on the ROC-COR tandem domains structure from *C. Tepidum* (Gotthardt, K. et al. 2008). Based on this model, the N- and C-terminal interacts, thus stabilizing the folding of the domain. While the COR domain is folded into 3 relatively independent globular parts, which suggests these 3 modules might remain stable if separated in short constructs. **B).** Limited Trypsin digestion on the full-length LRRK purified from HEK293T cells. The clearly defined bands were then cut out and test with Mass Spectrometry for peptide identification. **C)**. Diagram that summarizes information from multi-species alignment and secondary structure prediction. The helixes and strands are labeled red and yellow respectively, and the protease digested site from LP-MS are labeled with grey marks.

Table 2. List of ROC and COR domain constructs

Constructs	Vector and Tag	Expression test
1256-1540	pETDuet-1 and 6×his	expressed but insoluble
1256-1540	pGEX6p1 and GST	expressed but insoluble
1256-1540	pET28b and 6×his-sumo	expressed and soluble
1256-1561	pETDuet-1 and 6×his	expressed but insoluble
1256-1580	pETDuet-1 and 6×his	not expressed
1261-1540	pETDuet-1 and 6×his	not expressed
1261-1580	pETDuet-1 and 6×his	not expressed
1276-1520	pETDuet-1 and 6×his	not expressed
1276-1540	pETDuet-1 and 6×his	expressed but insoluble
1276-1587	pETDuet-1 and 6×his	expressed but insoluble
1301-1520	pETDuet-1 and 6×his	not expressed
1301-1540	pETDuet-1 and 6×his	expressed but insoluble
1301-1540	pETDuet-1 and 6×his	expressed but insoluble
1301-1540	pGEX6p1 and GST	expressed but insoluble
1301-1540	pET28b and 6×his-sumo	expressed and soluble
1301-1580	pETDuet-1 and 6×his	low expression
1301-1587	pETDuet-1 and 6×his	expressed but insoluble
1301-1587	pETDuet-1 and 6×his	expressed but insoluble
1307-1540	pETDuet-1 and 6×his	expressed but insoluble
1307-1561	pETDuet-1 and 6×his	expressed but insoluble
1307-1580	pETDuet-1 and 6×his	expressed but insoluble
1307-1643	pETDuet-1 and 6×his	expressed but insoluble
1312-1540	pETDuet-1 and 6×his	expressed and soluble
1312-1561	pETDuet-1 and 6×his	not expressed
1312-1580	pETDuet-1 and 6×his	expressed but insoluble
1312-1580	pETDuet-1 and 6×his	expressed and soluble
1312-1580	pET28b and 6×his-sumo	expressed but insoluble
1312-1643	pETDuet-1 and 6×his	expressed but insoluble
1312-1657	pETDuet-1 and 6×his	expressed but insoluble
1312-1672	pETDuet-1 and 6×his	expressed but insoluble
1312-1675	pETDuet-1 and 6×his	expressed but insoluble
1312-1832	pETDuet-1 and 6×his	expressed but insoluble
1315-1520	pETDuet-1 and 6×his	not expressed
1315-1529	pETDuet-1 and 6×his	expressed but insoluble
1315-1540	pETDuet-1 and 6×his	expressed but insoluble
1315-1587	pETDuet-1 and 6×his	expressed but insoluble
1320-1526	pETDuet-1 and 6×his	expressed and soluble
1320-1540	pETDuet-1 and 6×his	expressed but insoluble
1320-1561	pETDuet-1 and 6×his	expressed but insoluble
1320-1580	pETDuet-1 and 6×his	expressed but insoluble
1320-1580	pET28b and 6×his-sumo	expressed and soluble
1320-1587	pETDuet-1 and 6×his	expressed but insoluble

1320-1643	pETDuet-1 and 6×his	expressed and soluble
1320-1832	pETDuet-1 and 6×his	expressed and soluble
1323-1529	pETDuet-1 and 6×his	expressed but insoluble
1323-1540	pETDuet-1 and 6×his	expressed but insoluble
1323-1587	pETDuet-1 and 6×his	expressed but insoluble
1326-1540	pETDuet-1 and 6×his	expressed but insoluble
1326-1587	pETDuet-1 and 6×his	expressed but insoluble
1329-1520	pETDuet-1 and 6×his	expressed and soluble
1329-1529	pETDuet-1 and 6×his	expressed but insoluble
1329-1540	pETDuet-1 and 6×his	expressed and soluble
1329-1540	pGEX6p1 and GST	expressed but insoluble
1329-1540	pET28b and 6×his-sumo	expressed and soluble
1329-1561	pETDuet-1 and 6×his	expressed but insoluble
1329-1561	pGEX6p1 and GST	expressed but insoluble
1329-1580	pETDuet-1 and 6×his	expressed and soluble
1329-1580	pGEX6p1 and GST	expressed and soluble
1329-1580	pET28b and 6×his-sumo	expressed and soluble
1329-1587	pETDuet-1 and 6×his	expressed but insoluble
1329-1599	pETDuet-1 and 6×his	expressed but insoluble
1329-1618	pETDuet-1 and 6×his	expressed but insoluble
1329-1618	pGEX6p1 and GST	not expressed
1329-1643	pETDuet-1 and 6×his	expressed but insoluble
1329-1643	pGEX6p1 and GST	expressed but insoluble
1329-1655	pETDuet-1 and 6×his	expressed but insoluble
1329-1657	pETDuet-1 and 6×his	expressed but insoluble
1329-1832	pETDuet-1 and 6×his	expressed but insoluble
1329-1832	pGEX6p1 and GST	expressed but insoluble
1329-1845	pETDuet-1 and 6×his	expressed but insoluble
1329-1850	pETDuet-1 and 6×his	expressed but insoluble
1516-1618	pGEX6p1 and GST	expressed and soluble
1516-1618 (murine)	pGEX6p1 and GST	expressed and soluble
1599-1832	pETDuet-1 and 6×his	expressed but insoluble
1618-1832	pETDuet-1 and 6×his	not expressed
1618-1832	pGEX6p1 and GST	not expressed

Based on these strategies, total of 76 constructs were designed and tested (Table 2). These LRRK2 domain constructs were cloned into either pETDuet-1 (with N-terminal 6 × histidine tag for target protein), pET-sumo (with N-terminal 6 × histidine tag tandem with a sumo domain) or pGEX6p1 (N-terminal GST tag) expression vectors. They are transformed into Rosetta 2 (DE3) *E. coli* cells for protein expression. The expression test was conducted by comparing the samples before and after IPTG induction on SDS PAGE gels after separation of soluble and insoluble cell lysate. In this way, we were able to determine which constructs have robust soluble expression of the target protein (Figure A.28). After screening through these constructs, three constructs containing ROC domain were found highly stable and amendable for biochemical and biophysical experiments, which include ROC (1329-1520), ROC-h (1329-1540), ROC-cor1 (1329-1580). Among these three constructs, the ROC (1329-1520) had the highest yield and best stability. Thus, ROC (1329-1520) was used as the major working construct in the following biochemical and biophysical experiments.

B. GTPase activity and dimerization dynamics of ROC

ROC has dynamic dimer-monomer equilibrium in soltuion. Obtaining of a high yield and stable construct (ROC 1329-1520) enabled us to test the detailed biochemical and biophysical functions of the isolated ROC domain. The protein sample was purified by sequentially passing through a Nickel affinity column, followed by size exclusion chromatography. In the size exclusion purification step, the protein sample was eluted from the Superdex 200 column in two separate peaks (Figure 6A). Based on the calibration standard curve of the column, the first peak (at 93 ml) corresponded to a molecular mass of about 48 kDa, and the second peak (at 99 ml) of about 24 kDa. The theoretical mass of ROC is 23.5 kDa, thus the peak of 48 kDa and peak of 24 kDa is likely corresponding to the dimeric and monomeric conformation of ROC. Consistently, the protein sample of the two peaks resolved identically as a band of around 24 kDa on SDS PAGE, and both bands was successfully blotted with anti-LRRK2 antibodies on Western blots (Figure 6B). To further confirm the ROC oligomeric states in solution, we used multi-angle light scattering coupled with size exclusion chromatography (SEC-MALS) to determine the molecular mass of ROC (Figure 6C). The protein samples from the two peaks resolved in SEC-MALS as 46 ± 2 kDa and 25 ± 2 kDa respectively, thus consistent with the molecular mass of ROC dimer and monomer. These results confirm the existence of both dimeric and monomeric ROC in solution.

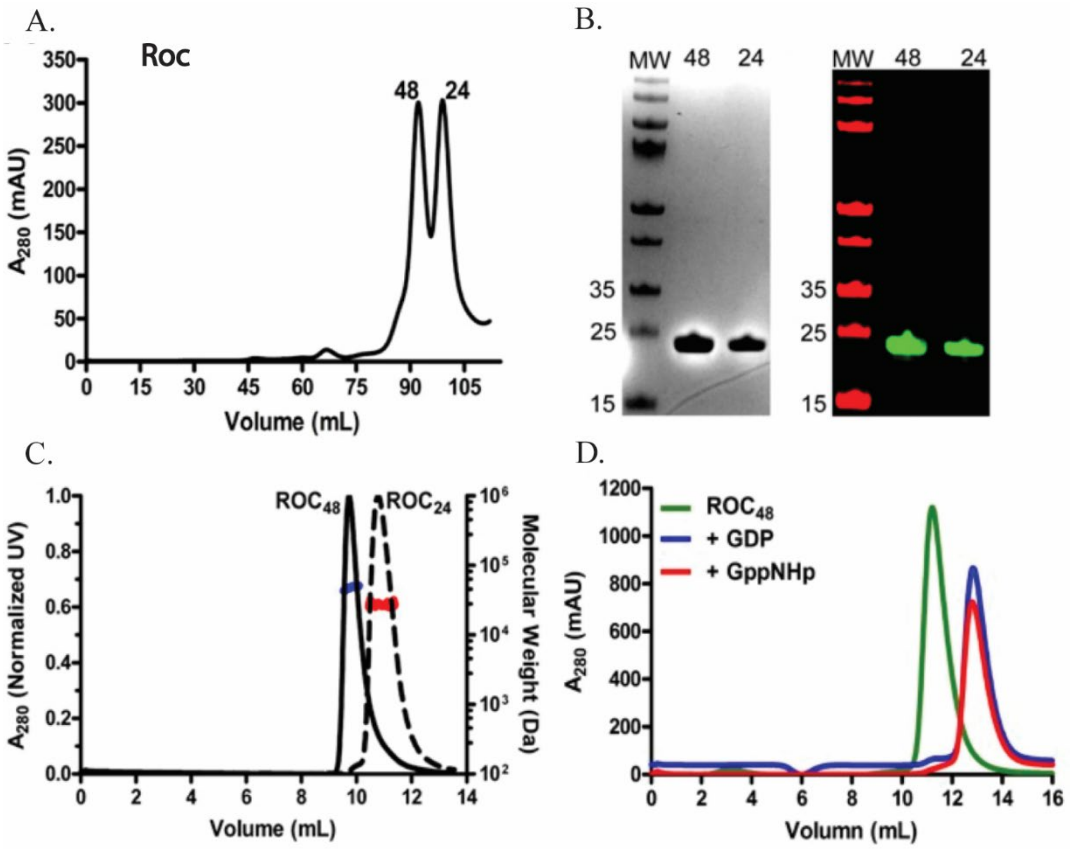


Figure 6. ROC exists in both dimeric and monomeric forms in solution
A). Elution profile of ROC in size exclusion chromatography (SEC) showing two peaks that correspond to ~48 and ~24 kDa. **B)**. SDS PAGE (left) of the ROC protein samples from the peaks of 48 and 24 kDa respectively. Western blot (right) with ROC antibody, thus revealing samples from both peaks are consisted of ROC. **C)**. SEC-MALS profile of the ROC from the two peaks, each plotted by the UV absorbance signal, and the calculated mass based on light scattering is marked blue and red for peaks of ~48 and 24 kDa respectively. **D)**. Both GDP/GppNHp loading with EDTA treatments convert the ROC dimer into monomeric conformation.

LRRK2 is a dimeric protein in solution as reported by multiple literatures (Greggio, Zambrano et al. 2008, Sen, Webber et al. 2009), thus suggesting that the oligomeric status as we observed on the isolated ROC is likely functional important. The existence of both dimeric and monomeric form of ROC in solution indicates that ROC may have a dynamic dimerization equilibrium, thus giving rise to the question that if the dimer-monomer equilibrium is associated with its GTPase function. To test this, we first used the standard nucleotide exchange protocol designed for small G proteins, which affiliates the nucleotide exchange rate by adding EDTA to strip off the Magnesium ion from the catalytic site during the incubation. Due to the high binding affinity of ROC to guanine nucleotides, the standard nucleotide exchange protocol uses high concentration of GDP or GppNHp (about 20 × ROC concentration) together with EDTA followed by 2 hours incubation in room temperature. After the EDTA affiliated nucleotide exchange, we analyzed the proportions of dimers and monomers using size-exclusion chromatography and found that almost all dimers were effectively converted to the monomeric form (Figure 6D). Thus, it suggests that the dimerization of ROC is dynamic, and the equilibrium of dimer and monomer can be altered by nucleotides binding. However, we observed in following up assays that the EDTA treatment alone without addition of guanine nucleotide also causes dimer-to-monomer conversion. Thus, the potentially different effects of GDP/GTP on the dimer-monomer equilibrium were likely overshadowed by the EDTA affiliated striping of magnesium ion.

The dimer-to-monomer conversion is efficiently induced by GTP binding. To have a comprehensive understanding on the effects of guanine nucleotides, we incubated the ROC dimer (93 % dimers) with different dosages of GDP or GTP respectively without

EDTA. The dimeric protein samples were titrated with different concentration of GDP or GTP ranging from 2.5 mM (4 × to ROC) to 10mM (15 × to ROC) and incubated for 6hrs in room temperature. Interestingly, we observed a GTP dosage-dependent dimer-to-monomer conversion, which equilibrated to 15 % dimeric and 85 % monomeric proportions respectively at 5 mM GTP (Figure 7A). In contrast, 5 mM GDP caused a smaller conversion that the dimeric and monomeric propositions are 71 % and 29 % respectively (Figure 7B). These data indicate that GTP binding leads to the dimer-to-monomer conversion, while GDP binding stabilized most ROC in dimeric form.

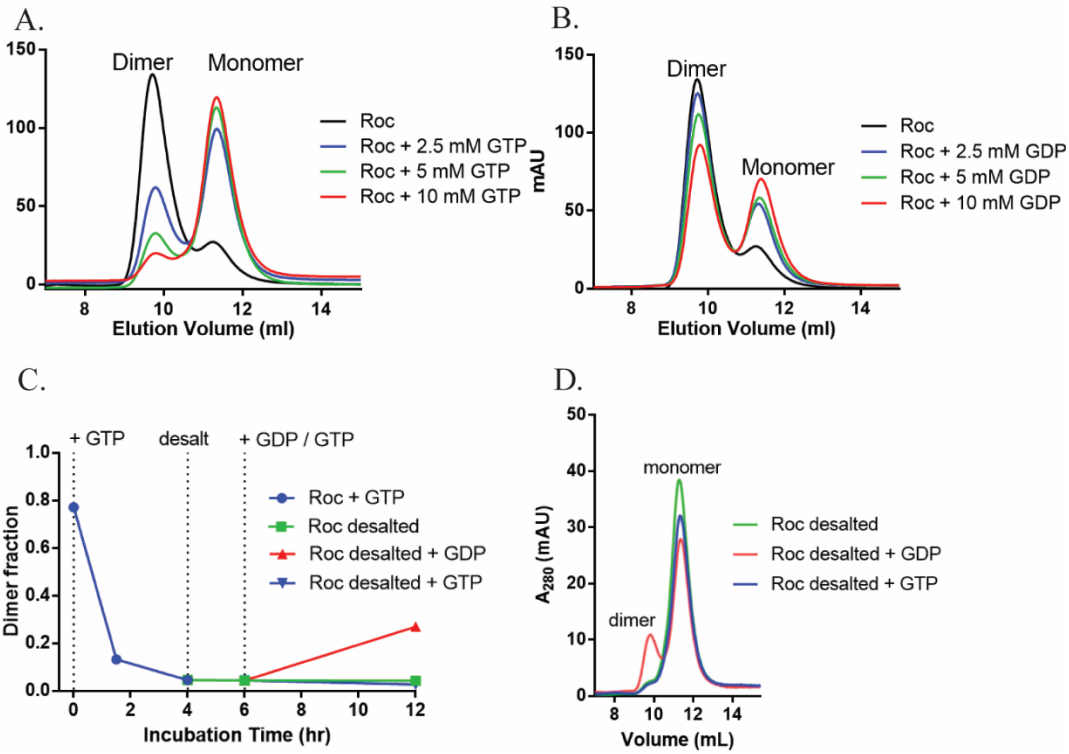


Figure 7. GDP and GTP drive the monomer-dimer interconversion of ROC
A) and **B)**. ROC dimer titrated and incubated with GTP and GDP respectively. Results show dramatic dimer to monomer conversion after GTP incubation, but most remain as dimer after GDP incubation**. C).** Diagram shows the dimer fraction of total ROC during the conversion cycle from dimer to monomer (with addition of GTP), then back to dimer (addition of GDP after desalting). **D)**. Elution profile of desalted ROC monomer incubated with either GDP or GTP. Result shows that the ROC monomer converts to dimer after GDP incubation.

To test how this dimer-to-monomer conversion is affected by GDP/GTP dosage and ratio, we titrated dimeric ROC (about 80 % dimeric) with different GDP/GTP ratios. Different ratio of GDP:GTP was prepared as 16:1, 4:1, 1:1, 1:4, 1:16 with total concentration of 16 mM (25 × to ROC) and incubated with the ROC sample in room temperature for 16 hours. The result shows that high GTP ratio leads to more monomeric (up to about 85 %) in the equilibrium, while higher GDP ratio leads to more dimeric (up to about 47 %). When GDP:GTP ratio is 1:1, about 35 % sample remained as dimeric, while 65 % was converted into monomers. (Figure 8C, 8D). Thus, the results consistently indicated that the GDP/GTP ratio regulates the dimer-to-monomer equilibrium, in which GTP binding induces the conversion but GDP binding stabilizes the dimeric form.

The monomer-to-dimer conversion is induced by GDP binding. The existence of dimeric ROC in solution suggests a potential monomer-to-dimer conversion also happens in solution. Given the stabilizing effect of GDP on the dimeric ROC, it is likely that GDP induce a monomer-to-dimer conversion in contrast to the effect of GTP. To investigate this, we developed an assay that started from a complete conversion of dimeric ROC into monomers by GTP, and then converted back to dimers by GDP. The dimer-monomer ratio was analyzed at each step of the process (Figure 9C, 9D). Under the "native" nucleotide exchange conditions without EDTA, the monomeric ROC protein sample (99 % monomeric) was prepared by incubating dimeric sample (85 % dimeric) with 16 mM GTP (25 × to ROC) for 4 hours. The monomeric ROC was then purified with desalt spin column to remove the extra GTP in solution. The desalted monomers were then split into three parts and incubated with GDP, GTP or without any nucleotide respectively for 6

hours. We found that after incubated with GDP for 6 hours, about 18 % of the ROC was reassembled into dimeric form, while GTP incubated ROC remains in the dimeric form. These results indicate that a reversible dynamic interconversion of ROC between dimeric and monomeric forms, in which GTP and GDP binding triggers the dimer-to-monomer and monomer-to-dimer conversions respectively.

To test how monomer-to-dimer conversion is affected by GTP/GDP ratio, we titrated different GDP/GTP ratios on the monomeric ROC (about 97 % monomeric) using the same method by GDP/GTP incubation and desalting as described previously. GDP:GTP ratio is also prepared as 16:1, 4:1, 1:1, 1:4, 1:16 with total concentration of 16

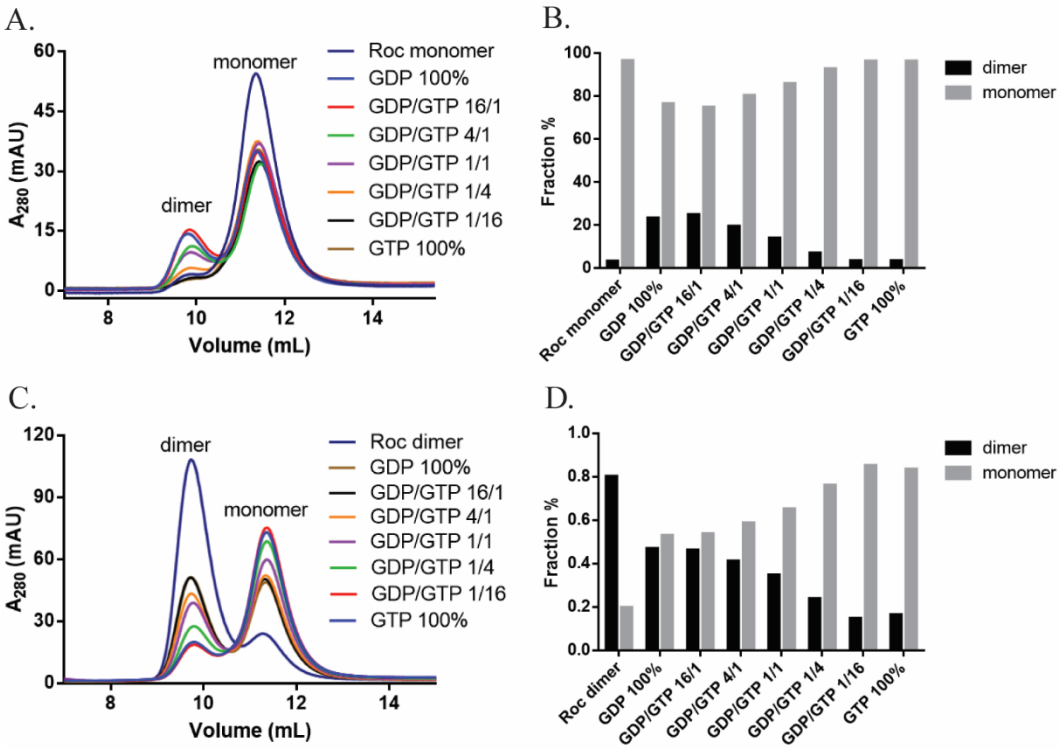


Figure 8. Dynamic equilibrium of ROC dimer-monomer modulated by GDP/GTP ratio in solution **A)**. Elution profile of ROC monomer titrated and incubated with different GDP/GTP ratio. The dimer and monomer fractions of ROC is calculated and showed in **B)**. Data showed the dimer-monomer equilibrium (converted from monomers) is correlated with the GTP/GDP ratio in solution. **C)**. Elution profile of ROC dimer titrated and incubated with different GDP/GTP ratio. The Calculated fraction of dimer and monomer is showed in **D)**. Data showed the dimer-monomer equilibrium (converted from dimers) is correlated with the GTP/GDP ratio in solution.

mM (25 × to ROC). After the 16 hours incubation, the monomeric ROC incubated with high GTP ratio was still monomeric (about 96 %), while the one with high GDP ratio had about 25 % converted into dimeric form. When GDP/GTP ratio is 1:1, about 18 % of the monomeric ROC is converted into dimeric form (Figure 8A, 8B). Thus, our data unambiguously demonstrate a GDP/GTP regulated reversible interconversion between dimers and monomers. In sum, GTP binding induces the dimer-to-monomer conversion, while GDP induces the monomer-to-dimer conversion, thus suggesting the dimer-monomer states are likely coupled with GTPase activity.

ROC is an active GTPase. To understand the GTPase activity in detail, we determined the steady-state kinetic parameters of ROC using the EnzCheck assay kit (Molecular Probes). The ROC protein sample were titrated with different GTP concentration, and the inorganic phosphate release from the hydrolysis of GTP was measured in a time course of 3 hours. The initial velocities of the GTPase reaction were then plotted as a function of GTP concentration. The plotted curve fits to the Michaelis-Menten equation, and the steady-state kinetic parameters were calculated accordingly: K_{cat} = 0.020 ± 0.001 min^{-1} and K_M = 553 ± 94 μM (Figure 9A). The low K_{cat} and high K_M values indicate that the isolated ROC is a slow enzyme.

G proteins commonly have a low intrinsic GTPase activity. The kinetics results of the ROC domain of LRRK2 is consistent with other G proteins. Ras, for example, has a similar low K_{cat} = 0.028 min^{-1}. It cycles between GTP-bound (active) and GDP-bound (inactive) states as a "molecular switch". After the GTP is hydrolyzed to GDP, Ras often remains in GDP-bound form due to its high binding affinity to GDP. The GDP-bound form of Ras cannot interact with downstream effectors, and thus it stays in the "inactive" state.

To speed up and regulate the exchange, guanine nucleotide exchange factors (GEF) affiliate the release of GDP and allow the binding of GTP (Vetter and Wittinghofer 2001). Due to its slow intrinsic hydrolysis activity, the G proteins often have a prolonged duration in GTP-bound form. The GTP-bound G proteins interact with effectors and trigger downstream signaling. Thus, it turns into the "active" state. This process is often regulated by GTPase activating proteins (GAP) to stimulate the hydrolysis activity. The observed low K_{cat} of ROC is in line with the activity of Ras family of GTPase, thus it might also require GEF and/or GAP in the cycling of GDP and GTP-bound forms. The Michaelis constant K_M is defined as the substrate concentration at which the reaction rate is at half-maximum. It is a measure of the substrate's affinity for the enzyme. In most fast enzyme reactions, the K_M often approximates the dissociation constant (K_d), which is defined as the substrate concentration at which half of the enzyme is bound with the substrate (or other ligands).

However, due to the slow intrinsic activity and tight binding affinity to the product to GTP and GDP respectively, we measured the K_d of ROC using a fluorescence polarization (FP) based binding assay. The non-hydrolysable GTP analog GTPγS and GDP were fluorescence labeled with BODIBY. The reason of using non-hydrolysable GTPγS instead of GTP is to avoid the complications resulted from GTP hydrolysis. Our result showed that K_d values for GTPγS and GDP are 4.1 ± 0.3 μM and 1.1 ± 0.1 μM respectively (Figure 9C and 9D). The binding affinity of GTPγS confirms that the Michaelis constant K_M does not represent the K_d of ROC. Although the K_d of ROC to its product GDP is lower compared to the K_d of GTPγS (thus suggesting a tighter binding to GDP than GTP),

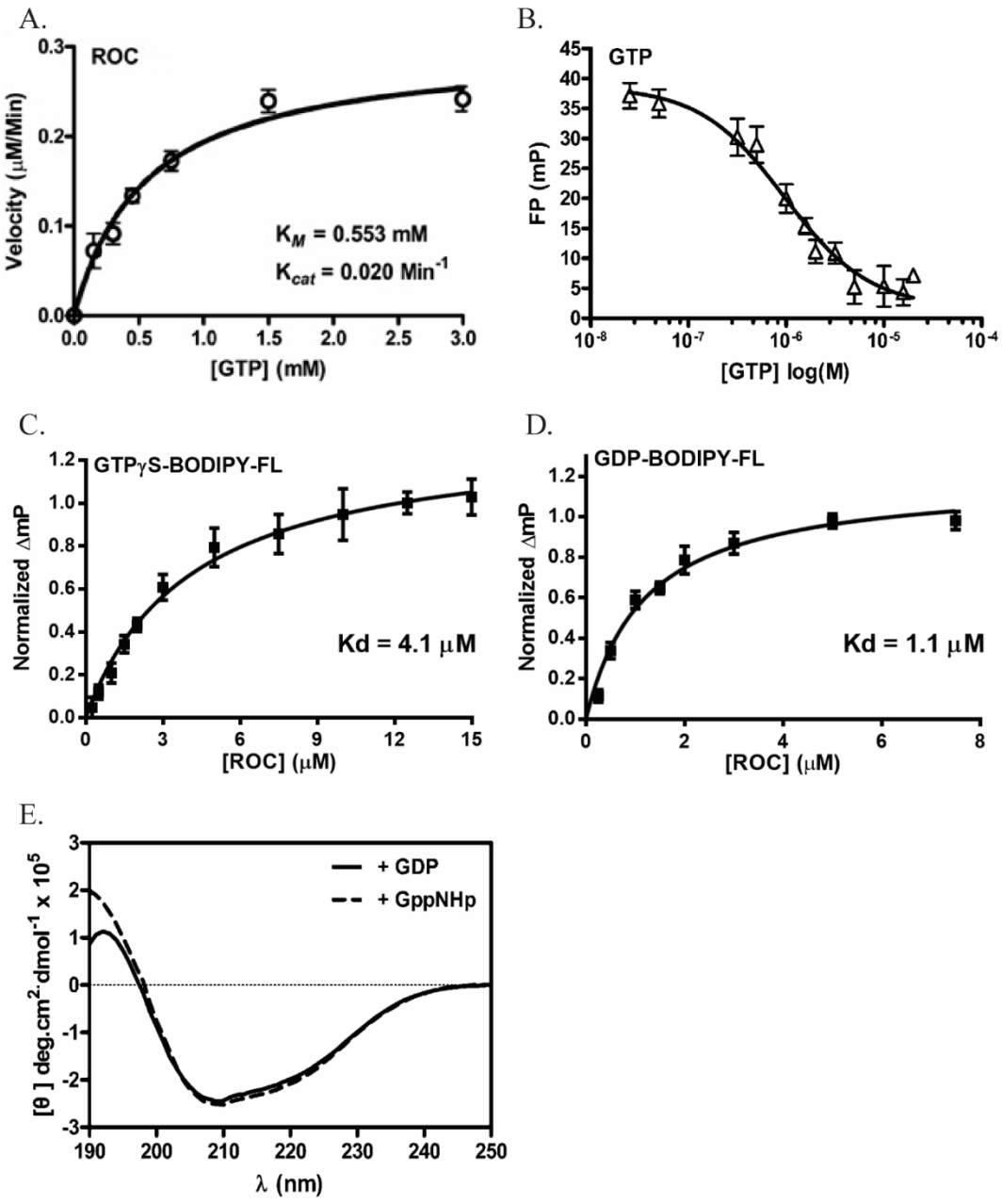


Figure 9. ROC is an active GTPase

A). Steady-state kinetics of ROC GTPase activity. **B).** Unlabeled GTP competition on GTPγS-BODIPY-FL that are pre-bound on ROC. Thus, the BODIPY labeled GTP analog bind to ROC in the same site as the natural GTP. **C)**. GTP binding affinity determined by FP assay with GTPγS-BODIPY-FL. **D).** GDP binding affinity determined by FP assay with GDP-BODIPY-FL. **E).** Circular Dichroism spectrum showing the ROC proteins bound with GDP or GTP analog (GppNHp) have no significant difference on secondary structure.

the micro-molar range K_d and high intracellular GTP concentration (physiological concentration of GTP are ~500 μM comparing to GDP of ~50 μM) suggests that ROC might not require guanine nucleotide exchange factor (GEF) for accessing its substrate, while the low K_{cat} of 0.02 turnover per minute suggests it might require an GTPase activating protein (GAP) for efficient hydrolysis of GTP.

C. Pathogenic mutations at R1441 reduce GTPase activity and disrupt dimerization dynamics

Residue R1441 in ROC is a hot spot of disease-associated mutations. Thus, the mutations may alter the GTPase function and disrupt the dimerization of the ROC domain. To study these pathogenic mutations, mutagenesis was conducted on the ROC construct (1329-1520). The protein samples of these mutants were expressed and purified using the same protocol as the wild type. The protein samples of R1441C, R1441G, R1441H all have a good protein yield (~1 mg purified protein from each liter culture).

Pathogenic mutations R1441C, R1441G, and R1441H reduce the GTPase activity. To understand the functional role of R1441, the GTPase activity of R1441C, R1441G and R1441H mutants are tested and compared side by side (Figure 10A). The results revealed that all mutants have reduced GTPase activity compared to the wild type. To compare the GTPase activity in detail, the steady-state kinetic parameters of each mutant were determined (Figure 11A, B, C). The K_{cat} values of R1441C, R1441G and R1441H are 0.003 min^{-1}, 0.005 min^{-1} and 0.009 min^{-1} respectively, comparing to the K_{cat} value of wild type 0.020 min^{-1}. While the Michaelis-Menten constant K_M of R1441C, R1441G and R1441H are 0.819 mM, 0.607 mM and 0.112 mM respectively, comparing to the wild type of 0.553 mM. Thus, the mutations on R1441 consensually cause reduced turn-over rate of the intrinsic activity for GTP hydrolysis, but the substrate access varies.

The R1441H leads increased GTP binding affinity, while R1441 mutations cause reduced affinity to GDP. To have a direct measurement and comparison of the GTP and GDP binding affinity, we then titrated the mutant proteins with BODIPY labeled GTPγS and GDP with protocol as described previously and calculate the K_d value based

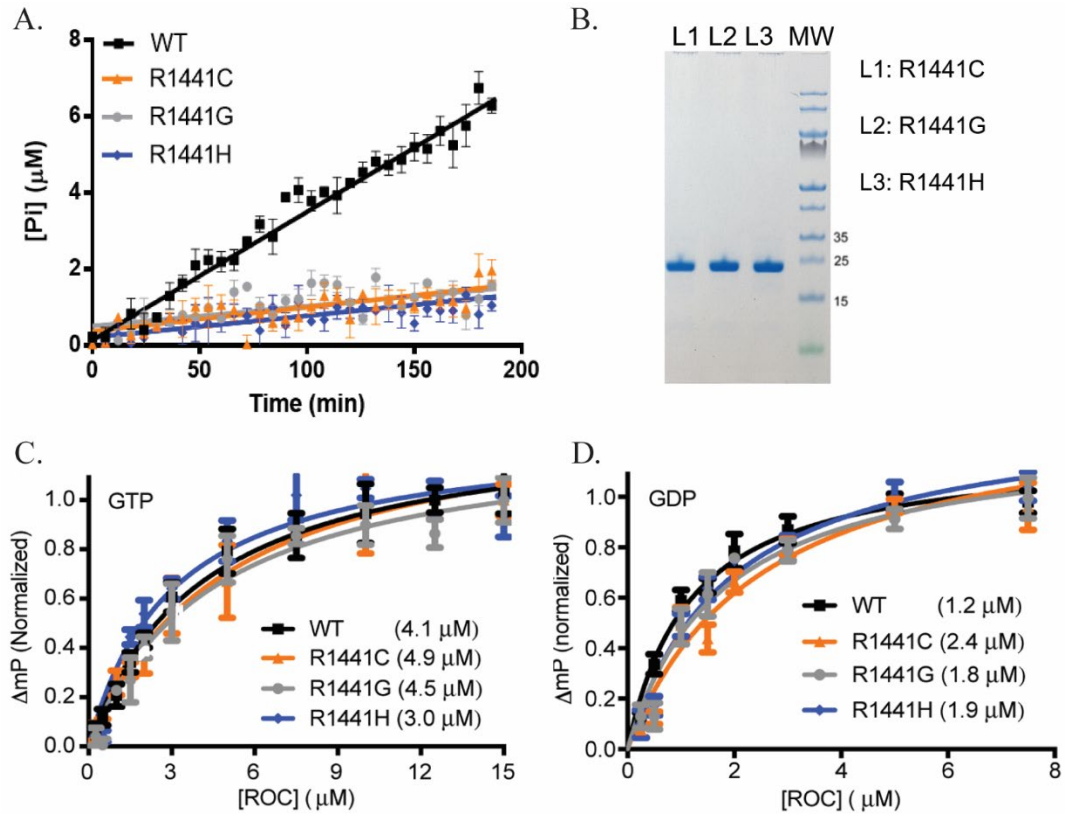


Figure 10. Pathogenic mutations R1441C, R1441G and R1441H in the ROC domain caused reduced GTPase activity

A) Side-by-side comparison of GTPase activity of the mutants with wild type ROC. All pathogenic mutants showed significantly reduced GTPase activity. **B)**. SDS PAGE of the purified protein samples of ROC mutants: R1441C, R1441G, R1441H, which were used in the GTPase activity assay. **C).** FP based GTP binding assay. The R1441H mutant of ROC showed decreased K_d of 3.0 μM, thus suggesting enhanced binding affinity, while the R1441C (K_d of 4.9 μM) and R1441G (K_d of 4.5 μM) have similar affinity as the wild type (K_d of 4.1 μM); **D).** FP based GDP binding affinity assay. All R1441C, R1441G and R1441H mutants showed slightly increase K_d of 2.4 μM, 1.8 μM and 1.9 μM respectively comparing to the wild type of 1.2 μM, thus suggesting reduced binding affinity to GDP.

on the fluorescence polarization result. The result shows R1441C, R1441G and R1441H has K_d value to GTP of 4.9 ± 0.8 μM, 4.5 ± 0.7 μM and 3.0 ± 0.4 μM respectively (Figure 10C), comparing to the 4.1 ± 0.3 μM K_d value of the wild type protein. While the values of K_d to GDP are 2.4 ± 0.4 μM, 1.8 ± 0.3 μM, 1.9 ± 0.3 μM respectively in R1441C, R1441G and R1441H (Figure 10D), comparing to 1.1 ± 0.1 μM K_d of the wild type. Thus,

the higher affinity for GTP observed in R1441H mutation is not a commonly shared property among the other disease-associated mutations (R1441C and R1441G). Thus, the alteration of guanine nucleotide binding affinity is likely not a key functional feature contributing to the reduced GTPase activity and development of disease.

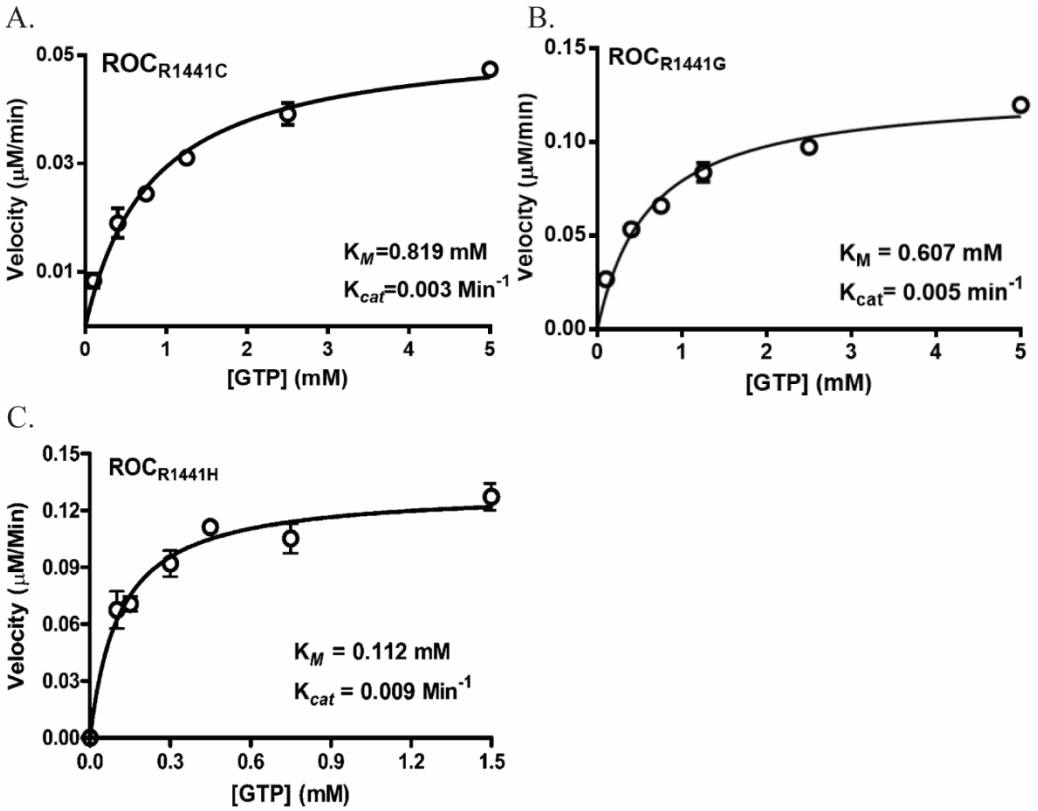


Figure 11. GTPase kinetics of ROC mutations R1441C, R1441G and R1441H
A). GTPase activity kinetics of the ROC R1441C mutant. **B).** GTPase activity kinetics of the ROC R1441G mutant. **C).** GTPase kinetics of the ROC domain R1441H. All pathogenic mutants showed reduced K_{cat} value, thus suggesting inhibited hydrolysis of GTP. However, the effects on substrate (GTP) accessibility varies: The R1441H showed reduced value of K_M, thus suggesting improved affinity to GTP; The R1441C showed increased K_M, thus the affinity to GTP decreased, while the R1441G mutant seemed to have little effect on the GTP accessibility.

To understand if the GTPase activity is disrupted by structural conformation and stability, we then tested if the R1441 mutations alter the secondary structure using circular dichroism (CD). CD is the absorption difference between the left-handed circular polarized light and right-handed circular polarized light. The circular dichroism spectrum in the wavelength range of 190 nm to 250 nm can be used to analyze the percentage of each secondary structure elements of the protein molecules in solution. The ROC samples of the WT, R1441C, R1441G and R1441H mutants were tested, and the CD spectra profiles indicated that the R1441 mutations have no significant differences comparing to that of the WT. Thus, it indicates there is no alteration on the secondary structure in these disease-associated mutants (Figure 12A). To determine if the R1441 mutations affect the thermal stability, we measured and compared the melting temperature (Tm) of R1441 mutants and wild type ROC protein (Figure 12B). The result indicates that R1441C and R1441G mutants have reduced Tm of 51.33 ± 0.14 and 50.13 ± 0.40°C respectively, comparing to the wild type Tm of 54.28 ± 0.01°C, thus suggesting these two mutations reduce the thermal stability of ROC. However, the R1441H mutation has the Tm of 53.75 ± 0.29°C, which is similar to the wild type, thus suggesting the thermal stability is not changed. Given the different effects of R1441 mutations, secondary structure and thermal stability are not likely the factors that cause consistently reduced GTPase activity.

Pathogenic mutations disrupt the dimeric conformation. Based on the crystal structure model, the exquisite interaction mediated by R1441 is critical for the dimerization of ROC. Thus, mutation on the R1441 residue is likely disrupt the dimeric conformation of the ROC domain. As expected, the R1441C, R1441G and R1441H were observed only eluted in a monomeric sized peak in Superdex 200 (Figure 12C). To confirm the molecular

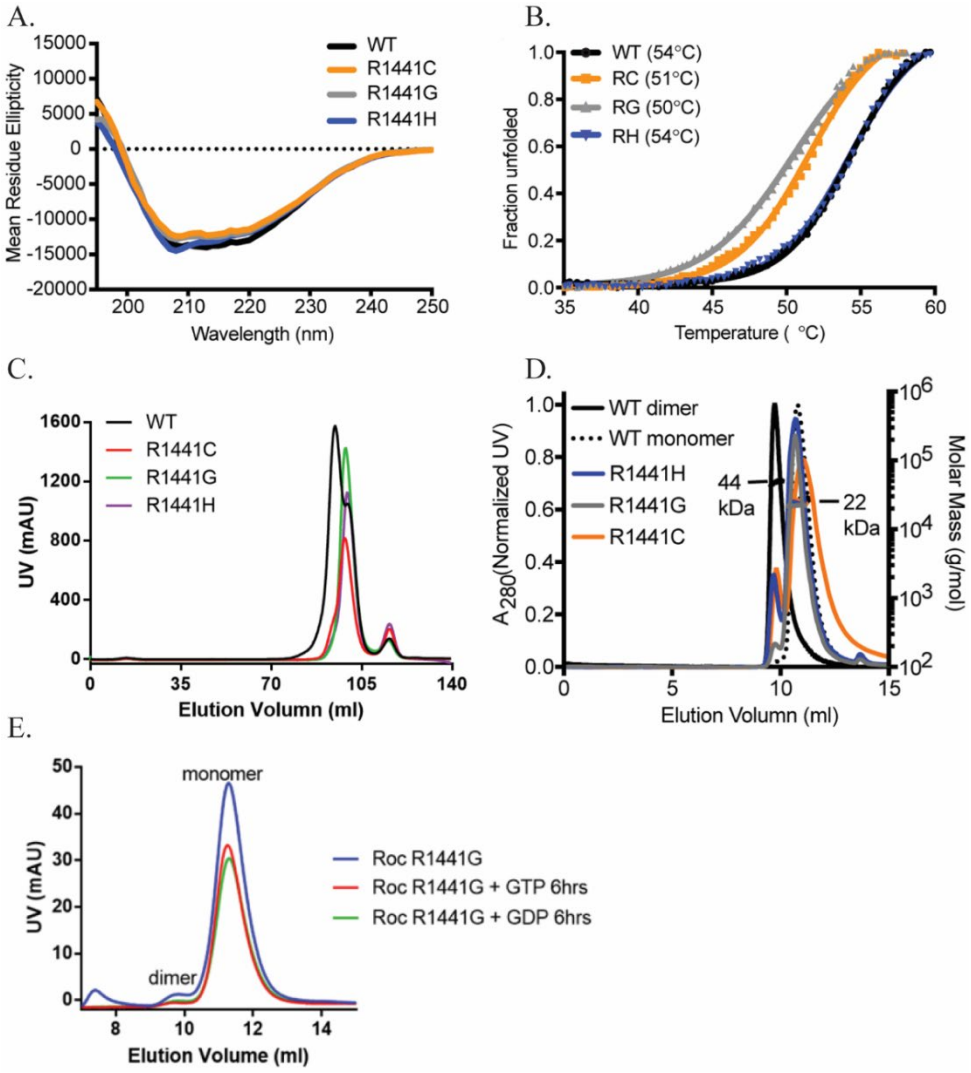


Figure 12. Pathogenic mutations R1441C, R1441G and R1441H in the ROC domain disrupt dimeric conformation

A) Side-by-side comparison of secondary structure conformation of ROC wild type, R1441C/G/H in solution by circular dichroism (CD). No significant alteration of secondary structure folding was observed. **B)**. Comparison of the thermal stability of the ROC wild type to the pathogenic mutants using thermofluor assay. R1441C and R1441G mutant showed shifted melt curve and decreased melting temperature, thus suggesting reduced thermal stability compared to wild type. However, the R1441H mutant protein showed no shifting of the melt curve, thus suggesting the thermal stability is not altered. **C).** Comparison of Superdex 200 size exclusion column elution profile. The wild type ROC protein showed mixed monomeric and dimeric forms, while the pathogenic mutants R1441C/G/H only has one protein peak corresponding to the monomeric form of the protein. **D).** Confirmation of the oligomeric state of the ROC mutants using SEC-MALS, which indicates the molecular weight of ROC oligos. **E).** R1441G stays as monomeric after GDP incubation, which converts monomeric wild type ROC to dimeric conformation.

mass of the protein molecules in the peak, the samples were tested using size exclusion chromatography coupled with multiple angle light scattering (SEC-MALS). Based on the results, all R1441 mutants are mainly in monomeric size (Figure 12D). The samples of R1441C and R1441H have shown a minor dimeric sized peak when eluted in the column with silicon-based matrix (Wyatt Tech) during the SEC-MALS analysis. Such dimeric peak was not observed during size exclusion chromatography conducted in the agarose-dextran based Superdex 200 or Superdex75. Thus, the minor dimeric peak is likely artificial in the silicon-based column. To test if the pathogenic mutants can still convert to dimeric form, the R1441G protein sample was incubated with GDP using the same protocol as the dimer-monomer interconversion experiments of the wild type. However, the results showed that the R1441G ROC stayed in the monomeric form (Figure 12E), thus suggesting the R1441G mutation disrupted the dimerization of the monomeric ROC induced by GDP.

To conclude, ROC showed a unique GTPase cycles coupled with dynamic dimer-monomer interconversion. However, the pathogenic mutations R1441G, R1441C, R1441H consensually disrupted the dynamic dimer-monomer interconversion of ROC in solution while reducing the hydrolysis of GTP, thus causing a prolonged GTP-bound monomeric form of the ROC domain. If the ROC domain functions as a molecular switch in a pathogenic signaling pathway in PD, the R1441 mutations cause a prolonged “ON” state of the signal.

D. Extended constructs of ROC suggest COR is likely involved in the dimer-monomer equilibrium

Using the ROC construct (1329-1520), we have observed an active GTPase domain that are coupled dynamic dimerization, which is disrupted by pathogenic mutations. However, two other ROC domain constructs, ROC-h (1329-1540) and ROC-cor1 (1329-1580), had shown some differences on their dimerization equilibrium.

ROC-h showed stabilized dimeric conformation. The construct ROC-h (1329-1540), based on the secondary structure prediction, contains one additional α-helix on its C terminal compared to the construct ROC (1329-1520). Different from the shorter ROC, the protein sample of ROC-h was eluted from the Superdex 200 column in only one peak (Figure 13A). SDS PAGE analysis showed the purified ROC-h has good purity (Figure 13B). To confirm the ROC-h oligomeric state, the sample was also tested by SEC-MALS as 54 ± 2 kDa (Figure 13C). The theoretical mass of ROC-h is 25.9 kDa, thus indicating a dimeric form of the protein. This is different from the observation of ROC (1329-1520), which has both dimeric and monomeric forms in solution (discussed in Chapter 1.3B). To test if the ROC-h dimer is dynamic, the sample was incubated with EDTA and GDP. The result showed a partial conversion of ROC-h into monomeric form (Figure 14A), thus suggesting a similar dynamic dimerization mechanism as observed in ROC. However, the C terminal extended α-helix seem to drive the equilibrium toward the dimeric side, thus suggesting it might be involved in the dimeric interface or crucial for the stability of the monomeric form.

ROC-cor1 is monomeric and has enhanced GTP binding. The ROC-cor1 (1329-1580), however, contains further extension of three additional α-helixes on the C

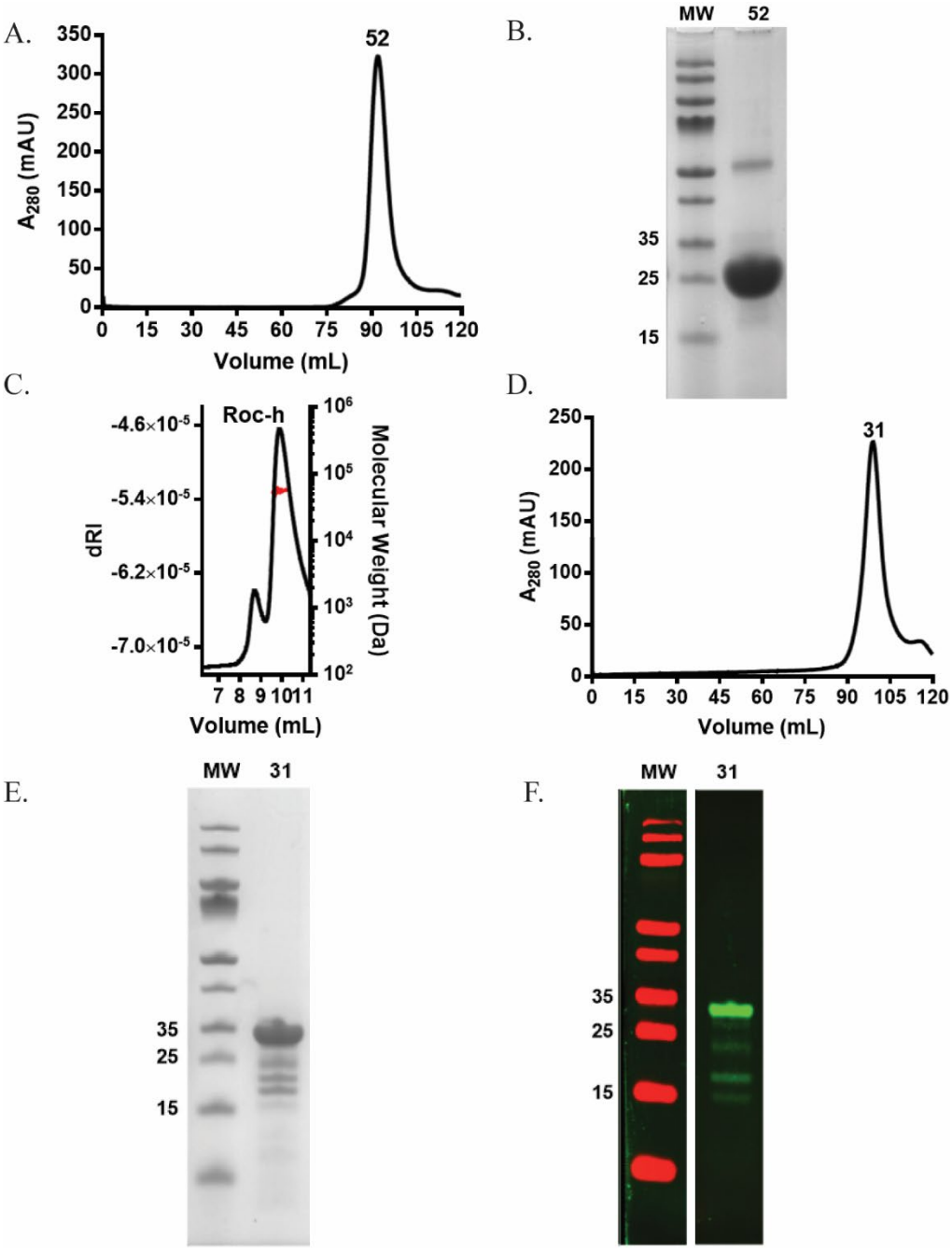


Figure 13. Extended ROC constructs: ROC-h (1329-1540) and ROC-cor1 (1329-1580) form mainly in dimeric and monomeric conformation respectively

A). Elution profile of ROC-h (1329-1540) in size exclusion chromatography (SEC) showing one peak that correspond to ~52 kDa. **B)**. SDS PAGE of the ROC-h protein samples from the peaks of 52 kDa respectively. **C)**. SEC-MALS profile of the ROC-h plotted by dRI (differential refractive index), the calculated mass based on light scattering is labeled in red. **D)**. SEC elution profile of ROC-cor1 (1329-1580) showing only one peak of ~31 kDa. **E)**. SDS-PAGE of the ROC-cor1 protein sample from the peak. **F)**. Western blot of the same ROC-cor1 sample, which reveals the minor bands that observed in SDS-PAGE are likely protease digested fragments of ROC-cor1.

terminal of ROC (1329-1520). According to the tertiary structure homology model based on ctROCO, these three alpha helixes interacts each other and form a 3-helix bundle. As the N-terminal region of the COR domain, the extended part (1521-1580) is a relatively independently folded motif in the predicted homology model of ROC-COR bidomain (Figure 5A). Interestingly, size exclusion chromatography showed only one peak after Superdex 200, which corresponds to a size of ~31 kDa (Figure 13D). The theoretical mass of ROC-cor1 is 30.5 kDa, thus suggesting it is monomeric in solution. The purified ROC-cor1 showed good purity on SDS PAGE (Figure 13E), however western blot has shown the lower bands are likely digested pieces of ROC-cor1 (Figure 13F), thus suggesting this construct has some flexible loops that are prone to protease digestion. To test if ROC-cor1 also has dynamic dimerization, the sample is desalted and incubated with GDP for 8 hours. However, ROC-cor1 stays monomeric before and after the incubation (Figure 14B). Thus, it suggests the extended part (1521-1580) on the C-terminal of ROC is likely stabilizing the monomeric conformation.

To test if the alteration on the oligomeric state of ROC-cor1 is associated with the GTPase function, we measured the GTPase kinetic and GDP/GTP binding affinities (Figure 14C.). Result of the GTPase activity shows dramatically reduced K_{cat} = 0.0015 min^{-1} (comparing to 0.028 min^{-1} of ROC), thus suggesting slowing of the GTP hydrolysis. However, the reduced K_M of 0.206 mM (comparing to 0.553 mM of ROC) suggests enhanced substrate GTP accessibility. Consistently, the binding assay results (Figure 14D) show K_d values of GDP and GTP binding are 0.8 and 2.4 μM respectively (comparing to 1.2 μM and 4.1 μM of ROC), thus suggesting an enhanced binding to GTP and GDP. Thus, the extended cor1 part (1521-1580) are likely stabilizing the GTP-bound monomeric

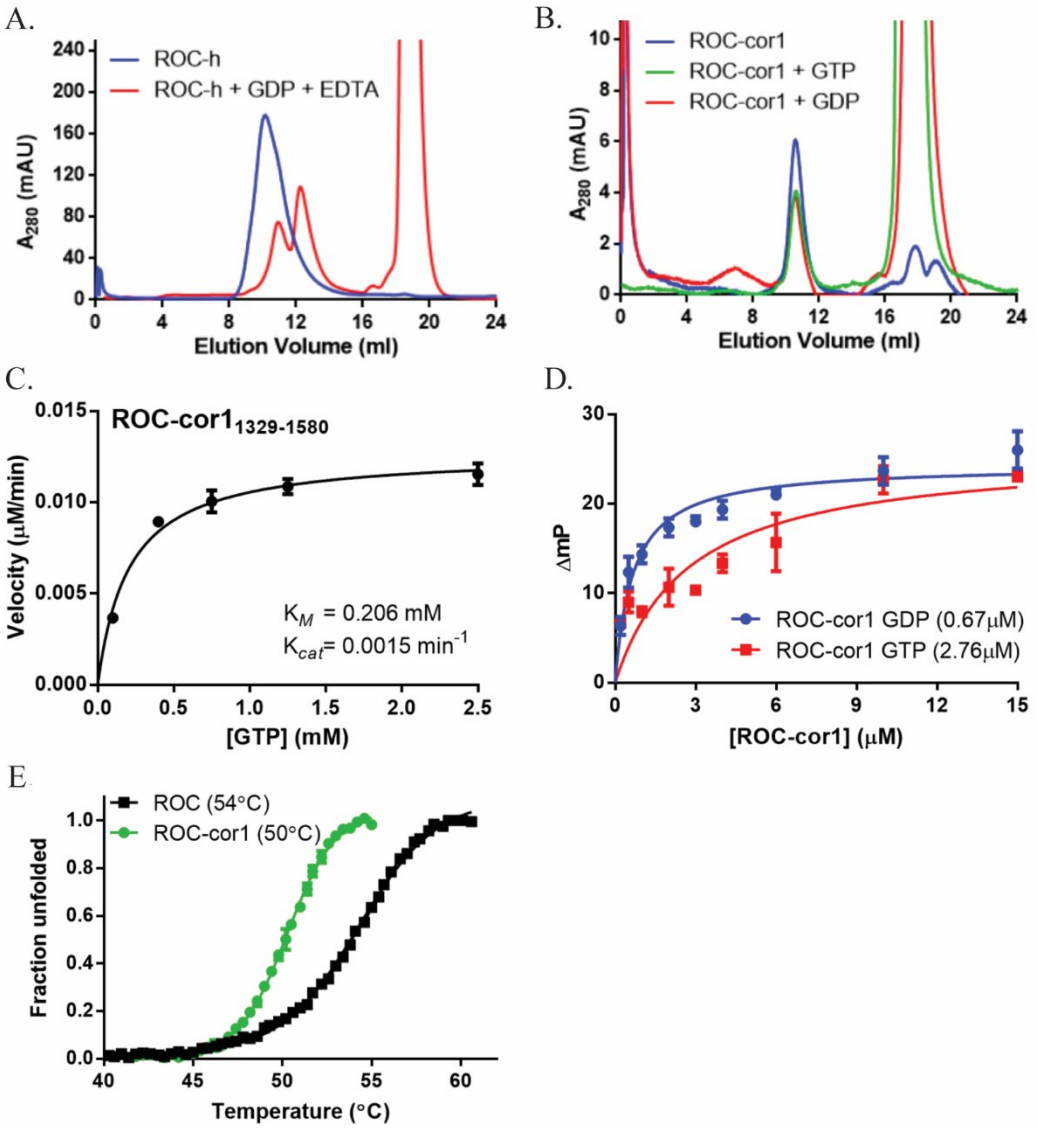


Figure 14. Dimer-monomer equilibrium and GTPase function of the extended ROC constructs: ROC-h and ROC-cor1

A). Dimer-to-monomer conversion of ROC-h by incubating the dimeric sample with GDP and EDTA. The result shows that ~60 % ROC-h dimers were converted to monomers. **B).** Modulation of the potential monomer-to-dimer conversion of ROC-cor1 by incubating the monomeric sample with either GDP or GTP. Similar to the R1441 pathogenic mutations, this ROC-cor1 construct cannot be converted from monomer-to-dimer. **C).** GTPase activity kinetics analysis on ROC-cor1 reveals enhanced substrate GTP accessibility, while dramatically reduce hydrolysis activity. **D).** FP based binding assay on GDP or GTP analogy (GTPγS) shows reduced K_d of GDP and GTP, thus suggesting tighter binding affinities. **E).** Thermal stability comparison between ROC and ROC-cor1. Result shows that ROC-cor1 has a reduced melting temperature of 50°C comparing to ROC of 54°C.

conformation. It lights up a potential regulatory mechanism of the COR domain that it is likely interact with ROC in the GTP-bound monomeric form.

However, it is still unclear how ROC-cor1 has slower hydrolysis than ROC. Also, we didn't find the existence of the dimeric form of this construct in solution, thus it is also unknown if it still has the dimer-monomer dynamic mechanism coupled to GDP/GTP binding cycles. Albeit the extended cor1 part on the C-terminal, the ROC-cor1 construct is less stable (Figure 14E) comparing to the shorter ROC (1329-1520). It is possible that ROC-cor1 has an unstable dimeric form, or the dimeric form only existed in a short time during the cycles. The dimeric form might facilitate the GDP release, while the stabilized monomeric ROC-cor1 has enhanced GDP binding thus causing reduced K_{cat} of GTPase activity.

Chapter 2. Structural basis for the function and conformational dynamics of ROC

2.1 Introduction

A. The controversial structures of ROC

So far, we have investigated the isolated ROC domain of LRRK2, which has both "Ras-like" and unique biochemical and biophysical characteristics. We observed a dynamic dimerization of ROC which is coupled with GTP/GDP exchange, in which the pathogenic mutations disrupt the GTPase activity and dimeric conformation in solution. Thus, it suggests that LRRK is likely has a unique structural mechanism different from the conventional Ras-like G protein switch mechanism.

Domain-swapped structure suggests a constitutive dimer of the ROC domain. A 2.0 Å crystal structure (PDB ID: 2ZEJ) of a ROC domain construct (1333-1516) were reported in 2008 (Deng, Lewis et al. 2008). In the structure model, the ROC domain forms a head-to-tail swapped dimer. Each ROC protomer folds into two lobes including a N-terminal "head" lobe and a C-terminal "tail" lobe which are linked by a distorted α-helix. The "head" and "tail" are exchanged between the two protomers, which form an extensive hydrophobic interactions and hydrogen bonds that bury about 5888 $Å^2$ of surface areas. A large portion of the interface is in the hydrophobic core of the dimeric protein. Thus, the domain-swapped dimer was suggested as a constitutive dimer serving as one functional unit. Two nucleotide binding sites were identified, which contains the P loop (interacting β-phosphate) and G4 loop (interacting guanine group) from the two protomer chains respectively, thus suggesting the dimeric conformation is required for guanine nucleotide binding (both GDP and GTP) and GTPase activity.

However, the literature of this structure model didn't provide direct evidence if this constitutive dimeric form indeed exists in solution. Instead the structure has been inconsistent with our biochemical and biophysical observations. We have observed a dynamic dimerization of the ROC domain (as discussed in Chapter 1), which is coupled with guanine nucleotide binding and the GTPase activity. The R1441 mutants, which are monomeric only, have reduced but still active GTPase activity. The binding affinities to GDP and GTP are also of similar level, thus suggesting the guanine nucleotide binding site are intact in the monomeric ROC. Thus, this structure model did not provide the insights to understand the conformational dynamics.

ROC is different from G proteins activated by nucleotide-dependent dimerization (GAD). Because the ROC domain has shown a guanine nucleotide binding associated dynamic dimerization, thus ROC might have a similar mechanism as G proteins activated by nucleotide-dependent dimerization. The G proteins activated by nucleotide-dependent dimerization (GAD) is a category of complex G proteins including signal recognition particle (SRP), dynamin and septins, which are regulated by homodimerization (Figure 15B) instead of the Ras-like GEF and GAP regulatory mechanism (Figure 15A). Due to the low intrinsic binding affinity for nucleotides (typically in the μM range), it does not require guanine nucleotide exchange factors (GEFs) to exchange GDP for GTP. Instead, the proteins are activated by nucleotide-dependent dimerization, in which the active site of one protomer is complemented and stabilized by the other protomer. This process is tightly coupled with the interaction with its effector. It comes to the "ON" state when they form a homodimers with GTP bound, while it goes back to "OFF" state after the hydrolysis and forms GDP bound monomers (Gasper, Meyer et al. 2009).

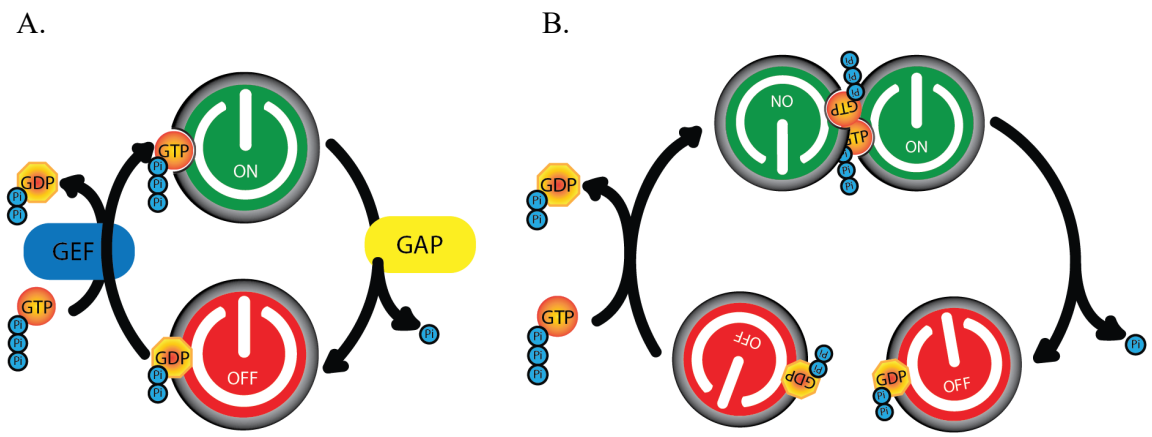


Figure 15. GTPase cycle of Ras family small G protein and G protein activated by nucleotide dependent dimerization (GAD)
A). The conventional mechanism of Ras-like G proteins, GTP loading allows the interaction with downstream effectors, thus turning the signal "ON". Due to the slow dissociation of GDP, the GTP loading is stimulated by guanine nucleotide exchange factors (GEFs) to release GDP from the active site. The hydrolysis of GTP to GDP is then accelerated by GTPase activating proteins (GAPs), thus turning the signal "OFF". This activating effect of GTP hydrolysis is through the interaction of GAPS with the active site of the Ras-like GTPase. **B).** The G proteins activated by nucleotide dependent dimerization (GADs) often has low affinity to nucleotides, thus does not require GEFs for the GDP dissociation. However, the hydrolysis of GTP is accelerated by a unique dimerization mechanism in which the active sites of both protomers interact and activate with each other, thus accelerates the hydrolysis of GTP.

However, the observed dynamic dimer-monomer interconversion of ROC is different from the mechanism of GAD. In contrast, the GTP-bound active ROC is monomeric in solution, while GDP binding drives the monomer-to-dimer conversion. Thus, ROC does not have a GTP-bound dimeric "ON" state. Instead, as suggested by the ROC-cor1 construct, the GTP-bound monomeric ROC is likely stabilized by interaction with COR *in cis*.

The ROC-COR bidomain structure of LRRK2 homolog ctROCO revealed a homodimer mediated by the COR domain. ROC-COR bidomain are conserved in ROCO protein family, thus suggesting LRRK2 and its homologs in the family are likely to share common functional and structural mechanism. As previously used in homology

modeling for constructs design, the structure of ctROCO from *C. tepidum* shed light on the potential structural mechanism of LRRK2.

A 2.9 Å crystal structure of the ROC-COR bidomain of ctROCO was reported in 2008. It is a 124 kDa complex protein of 1102 residues consisting of LRR, ROC and COR domains, thus shorter than LRRK2 that the N-terminal domains and kinase domain are absent. The structure of ctROCO revealed a unique COR domain mediated dimerization and GTPase activation mechanism (Gotthardt, Weyand et al. 2008). Based on the structure model, the dimeric interface is mainly mediated by the interactions between the C terminal regions of the COR domains, thus forming dimeric scaffold. The ROC domain intensively interacts with the *in cis* COR domain, thus holding the ROC domain in a position that the guanine nucleotide binding site is facing toward the other ROC domain on the other chain. However, only one ROC domain is found in the crystal structure of ROC-COR bidomain dimer. The shorter chain in the dimer is missing its ROC domain, which shows some electron density corresponding to the potential ROC domain but is not well resolved. Assuming the other ROC domain has a similar conformation as the observed one, an interface is likely to form between the catalytic sites of the two ROC domains, thus suggesting a dimerization-dependent mechanism for the GTPase function. An arginine residue, R543 outside of the conventional switch motifs, was found essential for the GTPase activity, of which the guanidinium group on side chain is likely pointing into the catalytic site of the other ROC in the dimer. Thus, a dimerization dependent "arginine finger"-like mechanism was proposed that the two ROC domains serving as each other's GAP.

However, the “arginine finger” R543 in ctROCO has no corresponding residue in LRRK2, while a corresponding pathogenic hotspot R1441 of LRRK2 is also absent in ctROCO. Thus, the mechanism proposed by the ctROCO might be different from LRRK2. Additionally, this COR domain mediated constitutive dimer was contradicted by a recent report that ctROCO also has a dynamic dimer-monomer conversion associated with GTP/GDP binding (Deyaert, Wauters et al. 2017). Thus, the structure model of the ctROCO is still not enough to explain the GTPase function and dimerization dynamics. Due to these reasons, we set the focus on determine the crystal structure of the human ROC (1329-1520) and ROC-cor1 (1329-1580).

B. Rationale and overview of this chapter

The biochemical and biophysical analyses in chapter 1 showed that the ROC domain has dynamic dimer-monomer interconversion coupled with its GTPase function. The pathogenic mutations at residue R1441 disrupt the dimeric conformation and cause reduced GTPase activity, thus suggesting the R1441 is likely associated in the dimeric interaction and GTPase function. As suggested by ROC-cor1 construct, the COR domain is likely interacting with ROC in GTP-bound monomeric form. The dynamic dimer-monomer interconversion of the ROC domain coupled with GDP/GTP binding state suggests the conformational change of the switch motifs is likely involved in the dimerization, while R1441 may play a role in the plasticity of the switch motifs. Thus, the overall goal of this chapter is to determine the structural mechanism of the biophysical and biochemical properties of ROC using X-ray crystallography.

To address these possibilities, we screened crystallization conditions for ROC and ROC-cor1, however only ROC was found with crystal hits. A 1.6 Å crystal structure of the ROC domain was solved, which contains well-resolved switch I and inter-switch loops missing in the previous reported crystal structure. The structure model shows a homodimer of ROC with a unique conformation that the switch motifs and inter-switch region are exchanged between the protomers (instead of a head-to-tail domain swapping). Comparing to conventional conformation of small G proteins, the two beta strands (the inter-switch) in between the switch I and switch II motifs are flipped out from the central beta sheet and inserted into the core of its counterpart in the dimer, thus acting as dual “latches” that fasten the two protomers together. The switch I and switch II motifs also adopt an unusual strained conformation like spring-loaded “hinges”, which mediate the dimeric interface while reach

out to the active site of the other protomer. Thus, the nucleotide binding site is composed of P-loop and G-binding motif from one protomer chain, while the switch motifs are from the other chain. Therefore, the GDP binding is likely independent to the dimeric conformation, while the γ-phosphate of GTP mediated interaction to the switch I and switch II motifs is likely associated with dimer-monomer interconversion. The dimeric interfaces are mainly formed by the inter-switch and switch II motifs, thus the conformation of switch motifs likely modulating the dimeric interaction of ROC. One of two major interfaces is mediated by the beta strand to beta strand hydrogen bonds between the inter-switch motif "latch" and the core beta sheet of the other chain, while the other interface is formed between the hydrophobic surface of the two switch II "hinge" motifs of the two protomers.

To verify if the unique swapped dimeric conformation indeed represents the ROC dimer in solution, a double cysteine mutant construct (ROC-dCys) is designed based on the structure model. If the dimeric conformation in crystal also exist in solution, it may form a disulfide bond cross the two protomers, thus covalently "locking" ROC in the dimeric form. Indeed, the purified dimeric ROC-dCys migrates as dimeric size in the non-reducing SDS PAGE, while shifts to monomeric size after treatment of reducing reagent DTT. Thus, ROC-dCys is indeed "locked" as dimer in soltuion. The ROC-dCys dimer has enhanced thermal stability when bound with GDP (effect reversed by addition of DTT), while the secondary structure folding remains the same as the wild type. However, the disulfide bond stabilized dimer of ROC-dCys has greatly reduced binding affinity to GTP, as well as dramatically decreased melting temperature comparing to the GDP-bound

sample or GTP bound wild type. Thus, it suggests the dissociation of the dimeric ROC is likely required for the loading of GTP.

The R1441 residue forms exquisite interactions with residues F1401 and T1404, which located between two distorted α-helixes (H2 and H3) in the switch II "hinge" motif. The guanidinium group of the R1441 side chain forms π-stacking with the phenyl ring of F1401, thus orienting two ω amines in the guanidinium group to form hydrogen bonds respectively with the carbonyl oxygen of F1401 and the hydroxyl oxygen of T1404. Thus, residue R1441 mediates an exquisite interaction with the switch II of the counter protomer, thus suggesting an essential role to stabilize the conformational strained switch II "hinge" in the dimeric ROC. GTP interacts with switch motifs and leads to their conformational change. Thus, the plasticity of switch II motif, which is modulated and stabilized by the R1441, is likely associated with GTPase activity. Consistently, the pathogenic mutations R1441C, R1441G and R1441H disrupt the interaction between residue 1441 to the switch II motif, thus altering the conformational plasticity of switch II and leading to impaired dimeric conformation and GTPase activity. Collectively, our data indicate that the hotspot of disease-associated mutations, R1441, directly interacts with the switch II motif of the counter protomer, thus modulating the dimeric equilibrium and GTPase activity.

2.2 Materials and Methods

A. Crystallization of ROC

To determine the molecular structure of the recombinant ROC (1329-1520) and ROC-cor1 (1329-1580), the protein samples were concentrated to 20 mg/ml and screened for crystallization conditions. The high throughput (HT) screening was done on automatic crystal screening robots: Phoenix (Art Robbins) and Oryx4 (Douglas Instruments). The crystallization screening was done on Intelli-3 96 well vapor diffusion plate (Art Robbins) and vapor batch plate (Douglas Instruments) with HT screening solution kits: Index, PEG/Ion, PEG-Rx, Crystal Screen (Hampton Research), Wizard Classic 1&2 (Rigaku) and PACT (Qiagen). The initial condition for crystalizing WT ROC was found in the condition containing 2 mM GDP, 100 mM KSCN, 25 % PEGMME, and 0.1 M Bis-Tris, pH 6.5. Some initial X-ray diffraction data set were collected using these crystals, and a partial model were built.

The optimal crystal morphology was achieved from condition of 3 µl of 20 mg/ml surface engineered ROC (K1460A and K1463A), mixed with 4 µl of 31 % PEG 3350, 0.25 mM NaCl, 1.3 M Bis-Tris, pH 6.5, 1 µl of the 1000 × diluted crystal seed. The crystal screening drop was then covered with 100 % paraffin oil, and observable crystals start to form after 2 days.

B. Surface residue engineering for crystallization

To optimize the crystal growth, surface engineering strategies were designed based on the partial model of WT ROC to improve the packing of the crystal, which give ride to K1460, K1463 and K1471 as potential mutation sites for surface entropy optimization. The Q5 site-directed mutagenesis kit (NEB) was used to generate the

mutated DNA of the ROC domain of LRRK2. The mutated plasmids were then transformed into Rosetta (DE3) cells for protein expression. The expression and purification protocols for surface engineered mutants are the same as described as for wild type and R1441 mutants in Chapter 2.

C. X-ray diffraction data collection and structure determination

The 1.6 Å X-ray diffraction data of the surface engineered ROC and a 1.9 Å wild type ROC was collected at 100K on the Pilatus3 6M detector (DECTRIS) at beamline 23-IDD of the Advanced Photon Source at the Argonne National Laboratory. The X-ray beam was set as wavelength of 1.03 Å and beam diameter of 0.2 μm. The crystals were cryo-cooled in liquid nitrogen after freshly harvested from the under-oil micro-batch crystallization. The diffraction data was then indexed and processed with software HKL2000 (Otwinowski and Minor 1997) in space group P21. The R_{merge} of the data set was 0.084. However, based on the data quality analysis by Xtriage in the program suite Phenix (Adams, Afonine et al. 2010), the data is twinned with a twin law of l, -k, h and twin fraction 0.44. Thus, the data was then either detwinned using the program detwin in the CCP4 suite before molecular replacement or refined with intensity based twin refinement in Refmac.

The PDB model 2ZEJ and a monomeric ROC homology model based on ctROCO (3DPU) were used as the search model to determine the phase using molecular replacement by Phaser in the Phenix programs suite. The model was then refined using Refmac (Winn, Murshudov et al. 2003) and Coot (Emsley and Cowtan 2004, Emsley, Lohkamp et al. 2010).

D. Twinning refinement

The data set was analyzed by Xtriage in program suit Phenix (Adams, Afonine et al. 2010) that it contains pseudo-merohedral twinning. The observed twin operator is l, -k, h, and twin fraction is 0.44. To detwin based on the calculated twin fraction and twin operator, the data was refined using intensity based twin refinement in Refmac (Murshudov, Skubak et al. 2011). Alternatively, the data set is processed by the program detwin in CCP4 program suit (Winn, Ballard et al. 2011). This program reads intensities from the original MTZ file and perform multiple twinning tests based on the provided twin operator. The calculation of the twin fraction in detwin is based on partial twin test plotted against theoretical expectations of cumulative intensity distribution, which is also used by twin analysis in Xtriage. To obtain an output MTZ file of detwinned data set, the twin fraction value of 0.44 was included in the input. The output MTZ file contains corrected intensities from a single twin fraction. It turns out that the Refmac twin refinement provide the best solution of twin refinement.

E. Expression and purification of disulfide bond stabilized ROC dimer

The double cysteine pairs for potential disulfide bond formation were designed based on the 1.6 Å structure model of ROC homodimer. The following pairs were made: R1398C + W1434C, T1343C + A1396C, T1348C + D1394C, V1340C + F1395C, Y1402C + N1437C, Y1402C + R1441C. Among the double cysteine mutants, R1398C + W1434C has the highest protein yield after purification, while the others have very low yield after purification. Thus, the R1398C + W1434C mutant was used for further testing on disulfide formation and dimer-monomer interconversion. The cloning method is the same as described for WT and other ROC mutants.

For optimal disulfide bond formation, the R1398C + W1434C mutant protein was expressed in the SHuffle T7 Express *lysY E. coli* (NEB), which contains disulfide bond isomerase (DsbC) and knock-out of thioredoxin reductases (trxB) and glutathione reductase (gor). The expression and purification were performed using same methods as for the WT ROC protein with buffers containing no DTT. For the non-reducing SDS PAGE, the protein buffer loading buffer were prepared as 100 mM Tris (pH 6.8), 4 % (w/v) SDS, 0.2 % (w/v) bromophenol blue, 20 % (v/v) glycerol, but without addition of DTT. The SDS gel preparation and electrophoresis were run as described in Chapter 1.

F. Molecular simulation

The structural simulation on the disulfide stabilized dimer of the ROC domain mutant R1398C + W1434C was performed using CHARMM (Brooks, Brooks et al. 2009). The program CHARMM is a molecular simulation program that containing sets of force fields for the analysis of molecular dynamics.

2.3 Results and Discussion

A. Determination of a 1.6 Å structure of ROC and twin refinement

To understand the biochemical mechanism and the GDP/GTP associated dynamic dimerization, we pursued to determine the crystal structure of ROC. To find a suitable condition for growing the ROC crystal, high throughput screening was done on the automated liquid handling robot Phoenix (Art Robins) with crystallization solution kits including: Index, PEG/Ion, PEG-Rx and Crystal screen (Hampton Research), Wizard Classic 1&2 (Rigaku Reagents), PACT suite (Qiagen). The ROC protein solution was concentrated to 20 mg/ml and mixed with 4mM guanine nucleotide (GDP/GTP/GppNHp) for the screening. The initial crystal condition was found on the index HT screening with condition of 2 mM GDP, 100 mM KSCN, 25 % PEGMME, and 0.1 M Bis-Tris, pH 6.5. The crystals observed take a thin plate cluster morphology, but we have managed to chop off some of the plate pieces as a single crystal for data collection (Figure 16A).

The initial X-ray diffraction data was collected based on these crystals to resolution of ~2 Å. However, the data collected lacks completeness and redundancy due to the bad morphology of these thin plate-shaped crystals. The stacking of these thin crystal plates further deteriorates the problem that the overlapping of reflection spots from multiple crystals made it difficult to index the space group and determine the unit cell parameters correctly. Thus, to solve the structure of ROC, the morphology of the crystals require optimization.

Optimization of ROC crystals by micro-seeding and surface residue engineering. To prepare single crystals, micro-seeding method was utilized with vapor

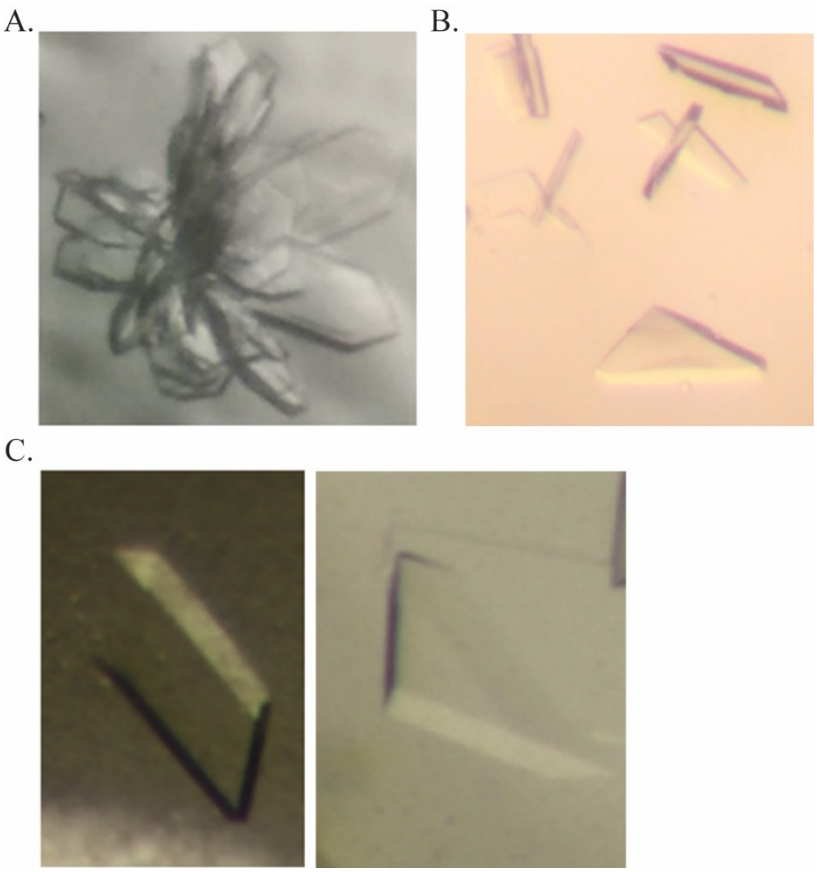


Figure 16. ROC crystals optimization
A). ROC crystals grow in the initial crystallization condition. **B)**. Improved ROC crystals using micro-batch setup with micro-seeding. **C).** Optimized ROC crystals with K1460A, K1463A surface mutations grow in micro-batch plates with micro-seeding.

diffusion and under-oil micro-batch crystallization setups. The crystals of ROC were crushed and resuspended in 30 % PEG. The crystal seeds were either prepared and used freshly or stored in -80°C after frozen in liquid nitrogen. Using diluted (1:1000) crystal seeds and micro-batch crystal growing system, although most crystals were still clustered, some single plate-shape crystals were formed and used for data collection (Figure 16B). These crystals diffract similarly to the initial thin cluster crystals to resolution of 1.8Å, and a nice data set was collected with 31422 unique reflections. The indexing, unit cell

refinement, integration and scaling were successfully processed in HKL3000. The space group is $P2_1$ (44.65 Å, 103.69 Å, 44.59 Å, 90.00°, 101.20°, 90.00°). The data set has 97.4 % completeness and R merge value of 0.072 (Table 3).

Because only a few thin plate crystals were able to collect complete data sets, surface residue engineering was performed to further improve the morphology of ROC crystals. The optimization was focused on lysine residues, which have flexible long side chain that are likely disrupting the molecular interaction during the packing of the protein molecules in crystal. The lysine residues K1460, K1463, K1471 were estimated as candidate mutation sites. These closely located lysine residues are not involved in crystal packing and show limited electron density for their side chains in the wild type structure, thus suggesting a negative effect for the crystal growth. Thus, ROC constructs with the lysine to alanine mutations were cloned and purified for crystal growth. We have found the combined K1460A and K1463A double mutation improves the crystal morphology to single thick plate shaped crystals (Figure 16C) in micro-batch crystalizing setups after seeded with diluted (1:1000) crystal seeds. The surface engineered K/A ROC (K1460A, K1463A) crystal is thick plate shaped, which diffracts to a higher resolution of 1.6 Å with unique reflections of 50680. The space group of the crystals belongs to $P2_1$ with unit cell of (44.634 Å, 101.883 Å, 44.61 Å, 90°, 100.953°, 90°). The data set has 94.5 % completeness and R merge value of 0.084 (Table 3).

Twinning issue of the crystals and twin refinement. However, further analysis conducted in Xtriage (included in Phenix) reveals that the both diffraction data sets of WT and surface K/A mutant (K1460A, K1463A) have twining issue. Twining occurs when a

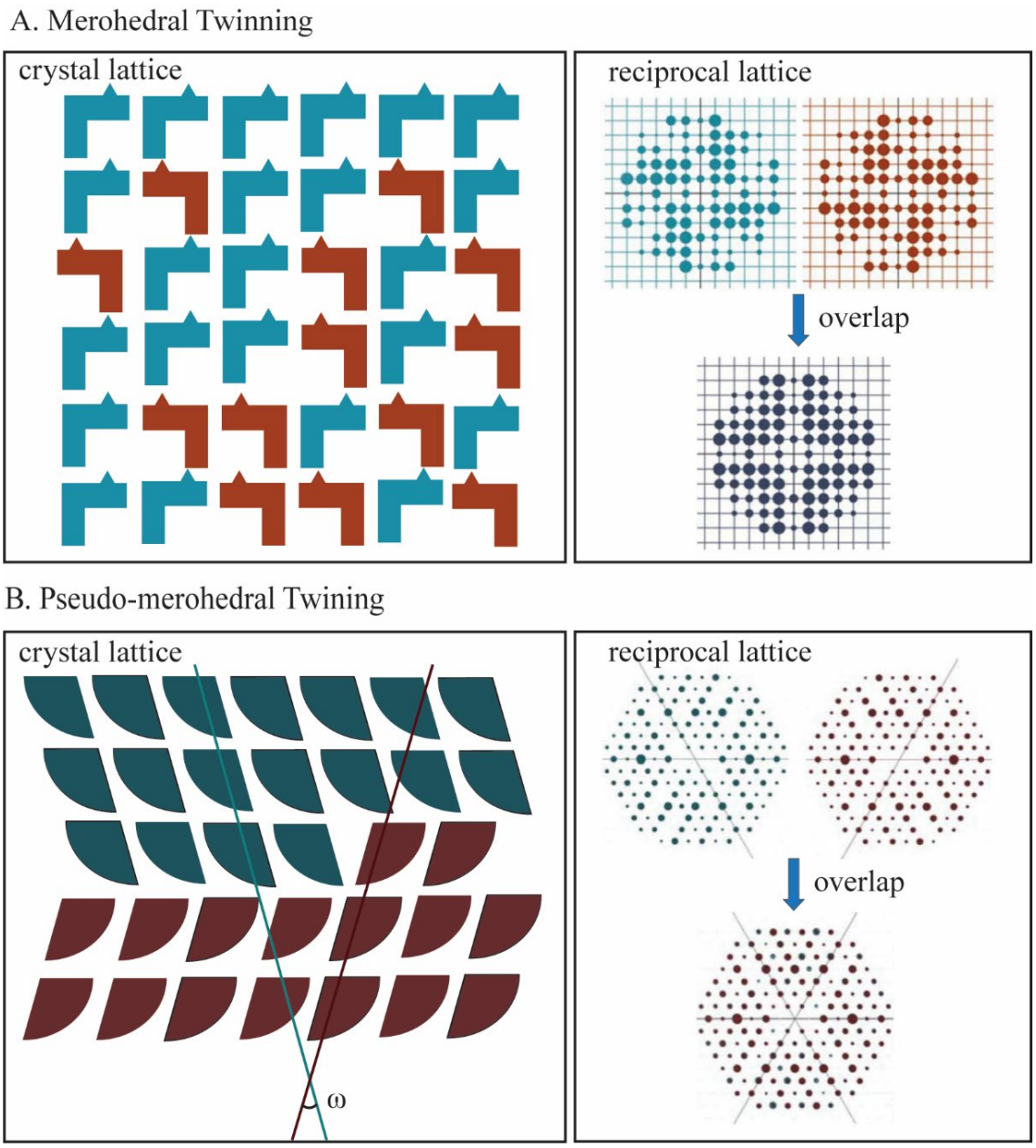


Figure 17. Crystal twinning and its effect on diffraction
A). Diagram showing merohedral twinning. The orientation of molecules in each domain are related by a symmetry operator and the crystal lattice is not changed. **B).** Diagram showing pseudo-merohedral twining. A pseudo-symmetry operator relates the two domains, but the crystal lattice is altered. The deviation angle between the lattices is termed as twin obliquity (ω). The smaller the value of obliquity, the more likely a pseudo-merohedral twining would happen.

crystal contains two or more discrete domains that oriented differently. Because both domains diffract X-ray, the observed diffraction frames contain overlapped reflection spots from both domains. Merohedral twinning occurs when the discrete domains in a crystal are related by a specific symmetry operator to the crystallographic axis, the diffraction pattern would overlap perfectly, and the observed data would look "normal" (Figure 17A). In merohedral twinning, the different oriented domain has the same crystal lattice as the original domain. However, the intensity of the reflection spots would have unusual intensity distribution. Thus, the shape of the cumulative intensity distribution is used to detect and analyze twinning, and thus the symmetric operator (termed as twinning operator) and the fraction of the discrete domain (termed as twinning fraction) can be calculated based on the intensity statistics.

In addition to merohedral twinning, pseudo-merohedral twinning is possible when unit cell can be transformed into a higher crystal system. This type of twinning only happens when the unit cells have specific parameters. In pseudo-merohedral twinning, the two (or more) domains are related by a pseudo-symmetry operator and the lattice between the domains are slightly deviated (Figure 17B). The deviation angle of the lattice planes between the discrete twinned domains is termed as twin obliquity (ω). The smaller the twin obliquity to close to 0 degree, the more likely the twinning would happen. For example, a monoclinic lattice with beta angle of near 90 degree is likely to form pseudo-twinning to a observed higher crystal system as orthorhombic. When the beta angle is close to 90 degree, the potential obliquity would be close to 0 degree.

Table 3. Statistics of X-ray diffraction data collection and refinement

Data collection			
Data set	ROC_{WT}	ROC_{KA}	ROC_{KA}-detwinned
Space group	$P2_1$	$P2_1$	
Unit cells	44.65 103.69 44.59	44.63 101.88 44.61	
	90.00 101.20 90.00	90.00 100.95 90.00	
Wavelength (Å)	1.033	1.033	
Resolution (Å)	43.80-1.88	43.82-1.59	
	(1.95-1.88)	(1.64-1.59)	
Completeness (%)	97.41 (82.63)	97.37 (78.93)	87.71 (47.64)
R_{merge}	0.072 (0.457)	0.084 (0.577)	
I/σ(I)	24.0 (2.3)	41.9 (1.7)	
Refinement			
Unique reflection	31422 (2692)	50680 (4069)	45088 (2454)
Protein atoms	2973	3029	3035
Solvent atoms	237	235	236
Ligands	60	60	60
R-factor (R_{free}) (%)	18.8 (24.7)	13.5 (15.8)	13.8 (17.4)
Average B-factor (Å)	46.97	41.31	44.4
R.M.S. deviations			
Bonds (Å)	0.019	0.023	0.032
Angles (°)	2.3	2.6	2.6
Ramachandran plot			
most favored regions (%)	95.07	89.19	90.8
Allowed regions (%)	3.84	7.3	6.5
Outlier regions (%)	1.1	3.51	2.7

To correctly process the data and solving the structure, we adopted two strategies to process the data sets: (1) process the twinned data together and apply intensity based twining refinement during the data refinement in Refmac; (2) calculate the twin operator and fraction, and "trim" the data using CCP4 software detwin to remove reflections from the twinned fraction. In the analysis of twin law by Xtriage, the data set was detected to contain pseudo-merohedral twinning with the observed twinning operator is l, -k, h, and twinning fraction is 0.44.

The refinement program Refmac is capable to perform intensity or amplitude based twin refinement in addition to the standard refinement (Murshudov, Skubak et al. 2011). The intensity based twin refinement uses reflection intensities directly, while the amplitude-based one uses intensities calculated from structure factor amplitudes. Theoretically, the intensity-based twin refine is better that it provides more accurate estimation on R-factor value. During the twin refinement, Refmac identifies 3 potential twin symmetry operators and calculated the Rmerge values of each operator, thus estimating the agreement between the twin-related refection sets. The smallest Rmerge suggests a good match of twin related reflection sets, and thus used for the refinement of each twin factors respectively.

The data sets for wild type ROC and K/A mutant were solved using molecular replacement (Phaser), in which a homology model of ROC built based on ctROCO was used as the searching model template. The data were then refined using Refmac with intensity based twin refinement. The Refmac corrects and refines both twin factors, thus utilizing reflections from both twin factors during the refinement. Using this method, the 1.88 Å ROC wild type data set (31422 unique reflections) was refined to R-factor of 18.8 % and R-free of 24.7 %, while the 1.59 Å K/A mutant data set (50680 unique reflections) was refined to R-factor and R-free of 13.5 % and 15.8 % respectively. Both the wild type and K/A surface engineered data sets reveal identical structure model, thus the data set of K/A mutant with better resolution and more unique reflections is mainly used for the following analysis and interpretations.

To confirm on the twin refinement solution, we also used program detwin (CCP4 programs) to process the data and correct the twining problem. The program detwin in

CCP4 is capable to (1) analyze twinned diffraction data and determine the twin operator and fractions, which can also be determined by program Xtriage in Phenix; (2) trim off twinned reflections from the original twinned data. Thus, it would provide a "de-twinned" data set, which is no longer twinned but contains less reflections (total reflections 45088). The completeness of the data is also reduced from 97.37 % to 87.71 %. Using the "de-twinned" data, the crystal structure was refined to similar R-factor and R-free of 13.8 % and 17.4 %. Comparing the two method for solving structure with twinning issue, both models provide electron-density map that unambiguously defined contiguous peptide chains of ROC. However, the "de-twinned" data has a significant reduction of total reflection and completeness, thus suggesting the twin refinement using Refmac, instead of detwin, is the better method to process twinned data.

B. ROC structure sheds light on the mechanism of the GTPase cycles coupled with dynamic dimerization

The electron density map defined two continuous peptide chains of ROC ranging from residues 1331 to 1518, which forms a homodimer in the asymmetric unit. The two protomers in the homodimer are related by a 2-fold symmetry with extensive interaction that buries 6461 $Å^2$ of surface area (Figure 18A). Two GDP molecules are bound at the guanine nucleotide binding sites of the dimer. The interaction is mainly mediated by the switch I, switch II and inter-switch regions, which are known flexible and key functional motifs involved in the dynamic conformational change between the active and inactive states upon GTP and GDP binding in Ras-like G proteins.

However, different from other Ras-like G proteins, the inter-switch motif consisting two β strands is unusually flipped open (Figure 22A, 22B) and inserted into the hydrophobic core of its dimeric partner *in trans*, thus wrapping up the dimer into a pretzel like conformation (Figure 18A). The interface between the two protomers are mainly composed of two regions: (1) hydrogen bonds interactions between the β3 strand in the inter-switch motif and β1' strand from its dimeric partner (Figure 18B), thus forming a beta sheet consisting of β2, β3 (inter-switch) and β1', β4', β5', β6' (hydrophobic core of its dimeric partner); (2) hydrophobic interaction between the switch II motifs, which is capped by dual hydrogen bonds with F1401' and T1404' by R1441 of its dimeric partner (Figure 18C). This unique "pretzel-like" conformation was previously also observed in a GDP-bound structure of rat Rab27b (Chavas, Torii et al. 2007). However, the functional meaning of this unique conformation was not addressed, and later on studies showed that

it folds into conventional Ras-like conformation when bound with its effector Slac2-a (Kukimoto-Niino, Sakamoto et al. 2008).

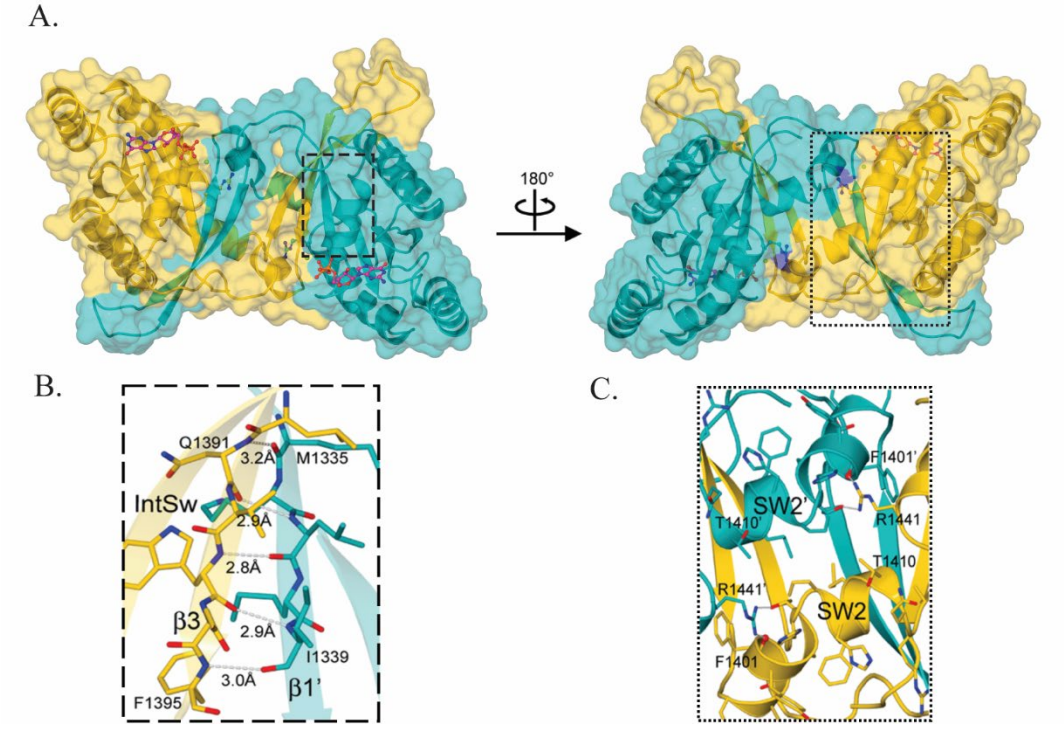
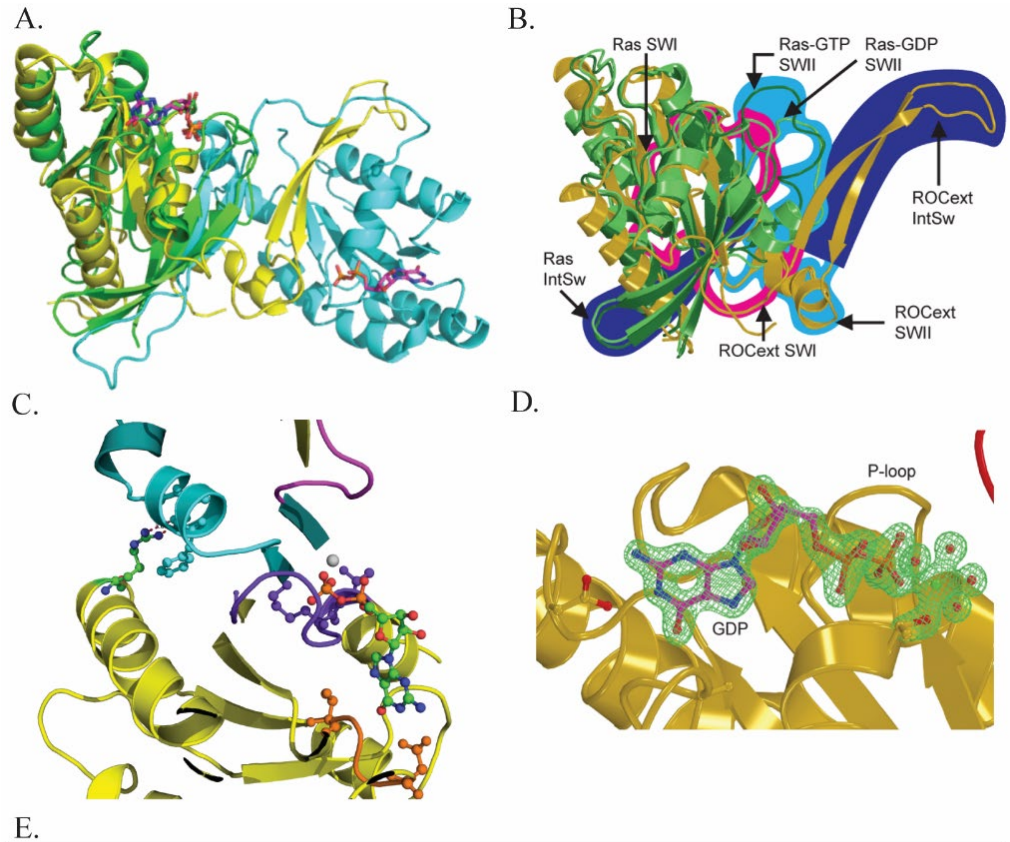


Figure 18. Crystal structure reveals a homodimer of ROC
A). Structure of ROC homodimer, showing extensive interaction mediated by exchanged switch motifs and inter-switch β hairpin. The two chains are colored as yellow and teal respectively. The dashed outline highlights the dimeric interfaces which are shown in B and C. **B)**. A zoomed view of the interface mediated by β strands from inter-switch and its dimeric partner. **C).** A zoomed view of the interface mediated by hydrophobic interaction between the two switch II motifs, which is capped by hydrogen bonds formed by R1441 and F1401, T1404 of the other chain in the dimer.

Comparing the tertiary structure of ROC dimer with Ras family G proteins suggests a very similar overall fold (Figure 19A), in which ROC dimer consists of all the structural and functional motifs including P-loop, Switch I, Switch II and G binding motifs. However, the catalytic sizes are formed by motifs from both protomers, and the switch I, switch II and inter-switch adopts a flipped-out conformation comparing the retracted inter-switch in Ras (Figure 19B). The essential motifs for GDP binding (P-loop and G binding

motif) are from the local chain of the binding site, while GTP interacting motifs (switch I, switch II) are from the other distal chain (Figure 19C). Two well-defined GDP molecules



E.

	G1 (P loop)	G2 (switch I)	G3 (Switch II)	G4
ROC-ROC dimer	1341 GNTGSGKT 1348	1356 KSDLGMQSAT 1369	1394 DFAGREEFYST 1404	1452 THLD 1455
human Ras	10 GAGGVGKS 17	29 VDEYDPT 35	57 DTAGQ E 62	116 NKCD 119
human Dynamin1	38 GGQSAGKS 45	56 FLPRGSGIVT 65	136 DLPGMTKVPVGDQPPDIE 153	205 TKLD 208
human GBP1	45 GLYRTGKS 52	65 FSLGSTVQSHT 75	97 DTEGLGDVEKGDNQN D 112	181 TLRD 184

Figure 19. ROC contains all structural motifs that are conserved in Ras G proteins
A). Superimpose of ROC dimer with GDP bound Ras (4q21). The overall folding is similar that the Ras aligned well with a ROC protomer. **B).** A close look at the alignment reveals that although all secondary structure elements aligned well, the switch I (SWI), switch II (SWII) and inter-switch (IntSw) motifs are remarkably different. The ROC protomer has a flipped and extended outward inter-switch with switch I and switch II as the hinge, while Ras has a retracted inter-switch that folds backward into its own core. **C).** A zoom in view of a catalytic site of the ROC dimer, which contains all characteristic structural motifs, including P-loop, switch I, switch II and G-binding. However, the switch I and switch II motifs are provided by the counter protomer, and R1441 residue from the local chain interacts with the counter switch II motif, thus likely stabilize its conformation. **D).** The GDP bound in the catalytic site has well defined electron density. **E).** A list of comparison of structure and functional motifs of ROC to other G proteins, which shows that ROC contains all the conserved residues for guanine nucleotide binding and GTPase activity.

are found in the catalytic sites of the ROC dimer, which interacts with P loop and G binding motifs similar to the interaction in Ras (Figure 19D).

The homodimer of ROC observed in crystal indeed exists in solution. As discussed in Chapter 1.1B, we have observed a dynamic dimerization equilibrium of ROC in soltuion. Using either GDP or GTP incubation, the ROC can be converted to either dimeric or monomeric forms. We used both dimeric and monomeric samples for crystallization trials. The GDP incubated dimeric sample forms crystals in about 3 to 5 days, while the GTP or GppNHp incubated samples more than 1 weeks to form crystals together with some precipitations. However, the result structures of both trails are dimeric, regardless of the starting conformation in crystallization setup, thus suggesting that the monomers of ROC may transit into dimers during the packing of crystal lattice.

To test if the dimeric conformation in crystal represents the dimers observed in solution, we designed double cystine mutations on residues, which are within ideal distance to form disulfide bond in trans between the two chains. Thus, if an inter-molecular disulfide bond is formed, it would covalently "lock" the ROC protein in the dimeric form. Accordingly, we designed serval pairs of double cystine mutations including R1398C + W1434C, T1343C + A1396C, T1348C + D1394C, V1340C + F1395C, Y1402C + N1437C. It turns out the R1398C + W1434C double mutant (Figure 20A) has the best stability and protein purification yield, while the other mutants are either not stable or insoluble. To express the protein with correctly formed the disulfide bonds, we used the SHuffle T7 Express *LysY* cells (NEB), which constitutively express a disulfide bond isomerase DsbC. The protein was then purified with similar procedure as the wild type in non-reducing buffer condition. The size exclusion elution profile showed two peaks (Figure 20B), which

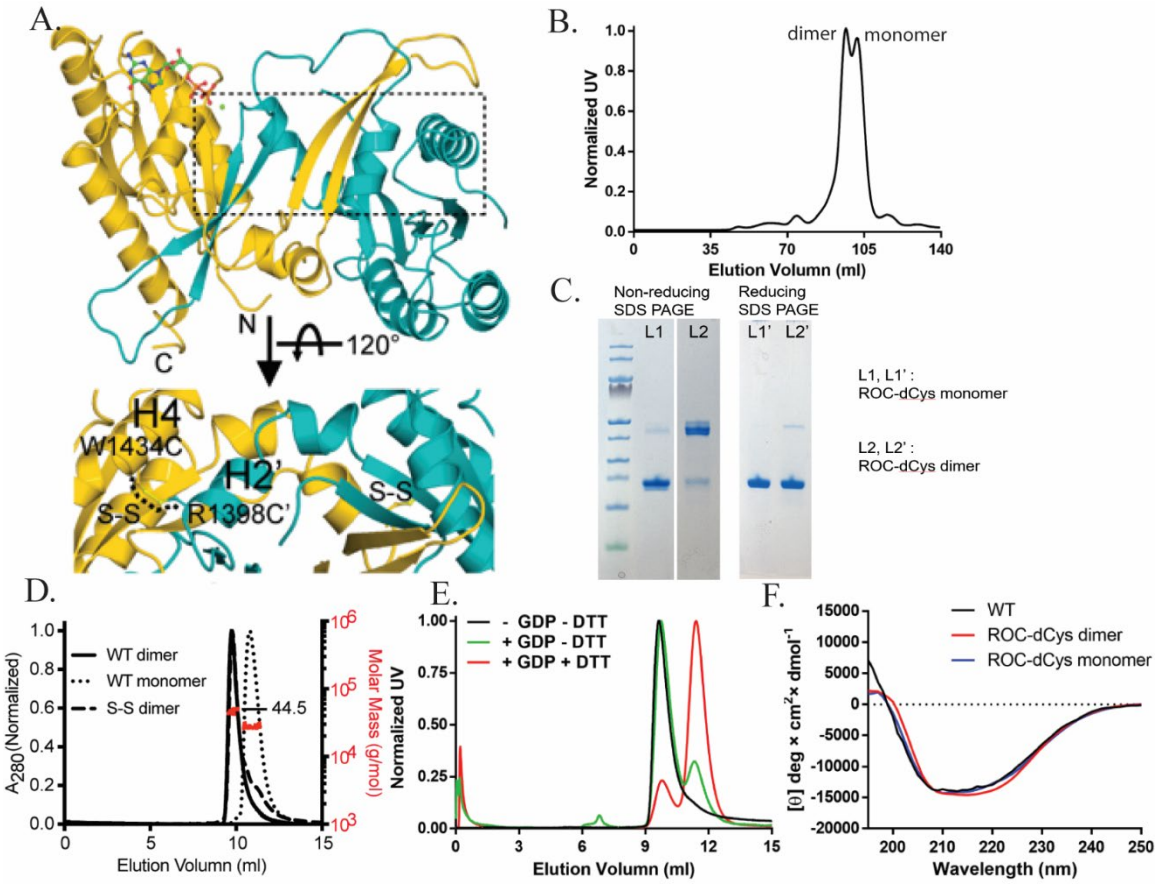


Figure 20. Disulfide bond stabilizes the dimeric form of ROC in solution
A). The double cysteine mutations pair (Roc-dCys) R1398C + W1434C designed based on the structure model. **B)**. Superdex 200 elution profile during the purification of the double cysteine mutant protein. It shows 2 peaks of ~48 and ~24 kDa, thus corresponding to dimeric and monomeric forms respectively. **C).** Reducing SDS-PAGE (right), both proteins migrate in the monomeric size. While in non-reducing SDS-PAGE (left), the dimeric sample migrate as dimeric size, thus suggesting the disulfide bond formed and covalently stabilized the dimer. **D).** SEC-MALS showed the molecular mass of disulfide bond stabilized Roc dimer is 44.5 kDa, thus confirming the it is indeed dimeric in solution. **E).** EDTA and GDP treatment on disulfide stabilized dimeric ROC, which can convert wild type ROC dimer into monomers completely. Data showed the disulfide stabilized dimer stays as dimeric, However, additional of DTT (reducing disulfide) together with EDTA and GDP converts into monomer, thus suggesting the disulfide bond stabilized the dimeric form. **F)**. Circular dichroism showing no significant secondary structure different between dimer, monomer and wild type dimer, thus suggesting no alteration on secondary structure.

correspond to the dimeric and monomeric forms respectively. The molecular weight of the disulfide bond stabilized dimer was confirmed using multiple angle light scattering (Figure 20D). The existence of both dimeric and monomeric peaks suggests the double cystine mutations alone causes no significant change on the equilibration of the oligomeric state of the ROC domain. Circular dichroism (CD) also showed that there is no significant difference on secondary structure between the R1398C + W1434C double mutation to the wild type ROC in solution (Figure 20F).

To analyze if the cysteines of R1398C + W1434C mutant ROC form inter-molecular disulfide bond that "locks" it in the dimeric form, non-reducing SDS PAGE and reducing SDS PAGE was conducted (Figure 20C). In SDS PAGE, the protein samples are denatured by heat shock and SDS detergent, thus it would migrate on the gel by the size of individual subunits instead of the native oligomeric complex. However, in the non-reducing SDS PAGE, the disulfide bonds covalently "lock" the two peptide chains of the dimeric protein together, thus it would still migrate as the size of the dimer. While in the reducing SDS PAGE, the addition of high concentration of DTT reduces and breaks disulfide bonds, thus the dimeric protein would migrate as monomeric size. Our result shows the R1398C + W1434C dimer sample migrates in a band of dimeric size on the non-reducing SDS PAGE, while it migrates as monomeric in the reducing SDS PAGE. Therefore, the inter-molecular disulfide bond forms in the dimeric sample of R1398C + W1434C double cysteine mutant, thus suggesting the dimeric structure model indeed represents the dimer ROC in solution.

We then tested if the inter-molecular disulfide bond changes the dimerization dynamics of the ROC protein. The ROC wild type protein converts from dimeric to

monomeric form after treatment of EDTA and GDP in room temperature. Therefore, the same treatment was conducted on the disulfide bond stabilized dimer sample of R1398C + W1434C (Figure 20E). The result showed that the sample remained as dimer after treatment of EDTA and GDP, but it was converted to monomeric form when treated together with DTT to reduce and break disulfide bond. Consistently, the disulfide stabilized dimer has increased melting temperature (Tm) of 61 °C than the wild type of 54 °C (Figure 21A), while DTT reduction reverse the melting temperature back to the wild type level of 54 °C (Figure 21B). Thus, the inter-molecular disulfide bond "locks" the ROC in a stable dimeric form.

Dimeric ROC is stable binding to GDP, while GTP binding may require monomeric conformation. With the disulfide bond "locked" dimer, now we can test how the dynamic dimer-monomer equilibrium could affects the function of ROC. As previously discussed on the dimer-monomer equilibrium, GTP can dramatically and efficiently converts wild type ROC dimers into monomeric form. Thus, the dimeric conformation might not be stable when bound with GTP. To test if alteration of the binding affinity to GTP is affected by the stability of GTP-bound dimer, melting temperature (Tm) of the disulfide bond stabilized dimer are measured with GDP or GTPgS (to prevent complication of GTP hydrolysis in higher temperature, the GTP analog is used instead of real GTP) in solution respectively. The result revealed that the Tm of the disulfide stabilized dimer dropped dramatically from 61 °C of GDP-bound to 52 °C of GTPgS-bound (Figure 21C), while the monomeric form of the double cysteine mutant has similar Tm of 55 °C and 52 °C when bound with GDP and GTPgS respectively. The results indicate that the GTP-bound ROC protein is more stable in the monomeric conformation, thus suggesting the

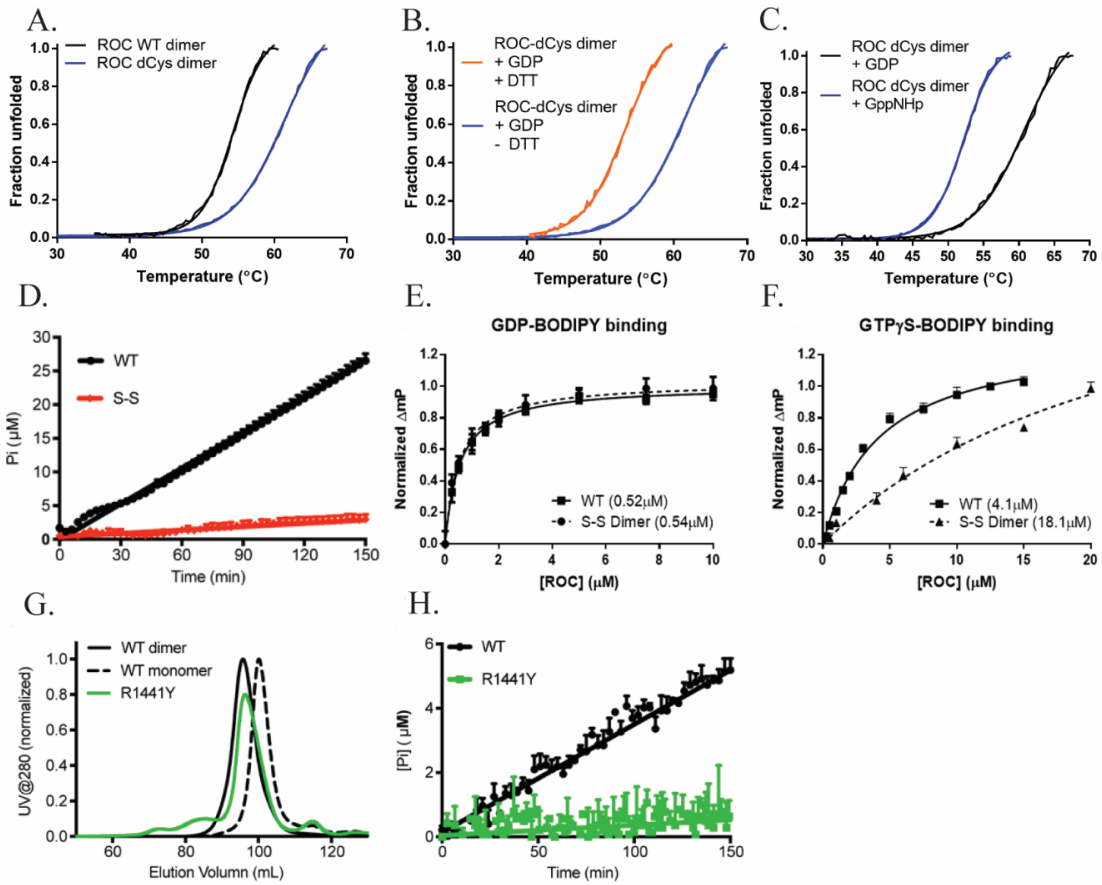


Figure 21. Disulfide bond stabilized ROC dimer showed reduced GTP binding affinity and GTPase activity

A). Thermofluor assay to compare the thermal stability of disulfide stabilized dimer to the wild type ROC dimer. The shifted melt curve reveals enhanced stability of the disulfide bond stabilized dimeric ROC. **B)**. Addition of DTT (to reduce the disulfide bonds) alters the melt curve of disulfide stabilized dimer back to similar level as wild type ROC dimer. **C).** Comparison of the thermal stability of disulfide stabilized dimer between GDP and GppNHp (GTP analog) incubated samples, which showed dramatically reduced melting temperature when incubated with GppNHp. Thus, it suggests the disulfide stabilized dimer are less stable when bound with GppNHp. **D).** GTPase activity comparison between wild type and disulfide bond stabilized dimer of ROC, in which the latter showed significantly reduced activity. **E).** Comparison between wild type and disulfide bond stabilized dimer on the binding affinity to GDP, which showed no significant difference. **F**). Comparison between wild type (albeit wild type ROC converts to monomeric form after binding of GTP) and disulfide bond stabilized dimer on the binding affinity to GTP, which showed significantly reduced GTP binding affinity of the disulfide bond stabilized dimer. **G)**. Superdex 200 elution profile of ROC with arbitrary mutation R1441Y, which showed only dimeric size protein peak during the size exclusion purification. Thus, the R1441Y mutated ROC is mainly in the dimeric conformation. **H).** GTPase activity comparison between wild type and R1441Y ROC, which showed dramatically reduced GTPase activity in the R1441Y mutant.

binding of GTP pushes the dimer-monomer equilibrium to the monomeric conformation. Consistently, fluorescence polarization based binding assay indicates the K_d to GDP is not changed in the disulfide bond stabilized dimer (Figure 21E), in which the K_d value is 0.54 μM comparing to the K_d of wild type ROC of 0.52 μM. However, the K_d value to GTP has dramatically increased to 18.1 μM, comparing to the wild type ROC of 4.1μM (Figure 21F). Thus, it suggests the "locked" dimeric form has dramatically reduced binding affinity to GTP. Thus, it suggests monomeric form is the preferred conformation for GTP binding. The GTPase activity assay showed that the disulfide stabilized dimer has dramtically reduced GTPase activity (Figure 21D). Therefore, the capability to form dimeric conformation is likely important for the GTPase function. Indeed, an arbitrary mutation R1441Y was also found to stabilize ROC in the dimeric form (Figure 21G), which has similar loss of GTPase activity (Figure 21H) as the disulfide bond stabilized dimer. Thus, it suggests the dimer-to-monomer conversion is likely also critical for the GTPase activity.

Comparison with a previous ROC structure reveals a likely error on chain assignment. A structure model of ROC was available in 2008 (2ZEJ), which showed a constitutive dimer of ROC model that is swapped head to tail (Deng, Lewis et al. 2008). Based on this model, the formation of the dimer is essential for both the stability in solution and GTP/GDP binding capability, however it is not consistent with what we observed in solution that ROC is an active GTPase in monomeric form and has dynamic dimer-monomer interconversion. Thus, we compared the two structure to investigate the differences.

Superposition of the two dimeric structures shows that they are nearly identical (RMSD 0.3 $Å^2$), however there are major differences on the chain assignment on β strand

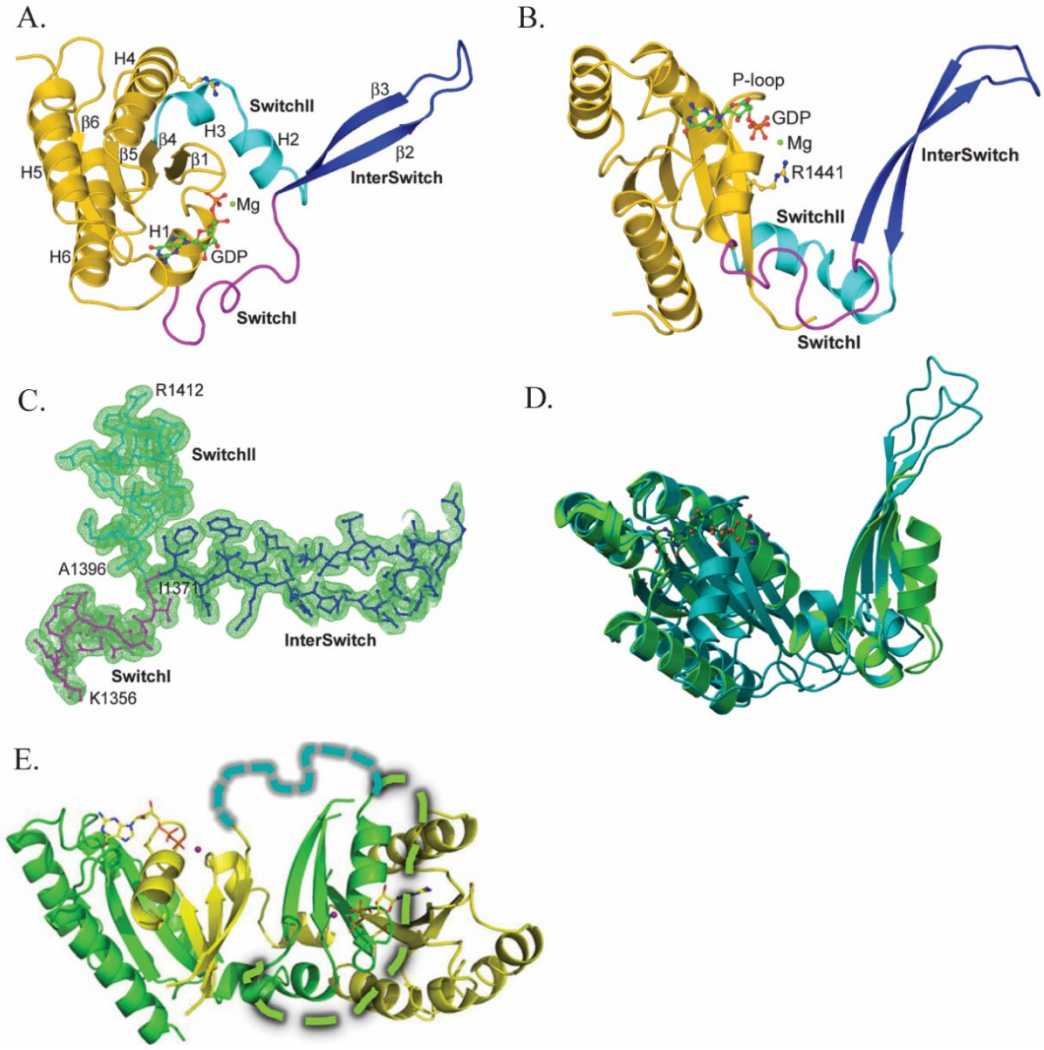


Figure 22. Unique conformation of the inter-switch and switch motifs suggests a potential "latch" and loaded "hinge" mechanism for the GTPase coupled dynamic dimerization of ROC
A). & **B).** Display of one protomer of the dimeric complex, the inter-switch motif is flipped and extended outward, which inserts into the core of the other protomer in the dimeric conformation. Thus, the inter-switch motif of each protomer acts like a "latch" that extends out and locks both protomers together. The switch I and switch II motifs are structurally strained, thus likely acting as spring loaded "hinges", which might drive vast conformational change when the dimer dissociates. **C).** Electron density map of the switch I, inter-switch and switch II motifs (σ=2.0). **D).** Protomer structural alignment of our ROC model to the previously reported ROC model (2ZEJ). The structural position of most atoms are almost the same, except a main difference on the chain assignment of β-strand 1 (β1) and α-helix 1 (H1). E). The previous reported ROC model (2ZEJ) was missing the loop (shown as teal dash line) around switch I motif, which is well-defined in our model, as shown by electron density map in **C).** Thus, the previous model likely made an error on the chain assignment, which assumes the missing loops retracts backward to guanine nucleotide binding site (shown as green dash line).

1 (β1) and α helix 1 (H1) (Figure 22D). In the model 2ZEJ, a large switch I-containing loop between H1 and β2 is missing, thus the assignment on the N-terminal fragment became ambiguous. Thus, the previous model has interpreted the model as the entire N-terminal part before switch II motif are all flipped and extended outward, including β1, P-loop, H1, switch I loop, and inter-switch (β2 and β3). Thus, albeit the 2ZEJ has precisely the same locations in terms of atomic interpretation of the electron density map, there are systematic differences on how the switch I loop would connect the N and C terminal parts in a peptide chain. According to 2ZEJ, the P-loop and G binding motifs in each catalytic site are separately from chain A and chain B respectively, thus indicating a dimeric dependent binding for GDP. However, we have in contrast observed both monomeric mutant (R1441C/G/H) and dimeric ROC (stabilized by disulfide bond) have the same binding affinity to GDP. Also, our electron density map has continuously defined the previously missing switch I loop (Figure 22C), which unambiguously reveals that only the inter-switch motif (β2 and β3) is flipped outward and exchanged between the protomers in dimer. Thus, the β1, P-loop, H1 part in model 2ZEJ was likely mis-assigned and swapped to the opposite chain (Figure 22E).

ROC structure sheds light on the mechanism of the dynamic dimerization coupled with GTP/GDP binding. The Ras-like G proteins are molecular switches that toggle the switch motifs between "ON" and "OFF" conformation upon binding to GTP or GDP respectively. The switch I and switch II motifs interact with the γ-phosphate of GTP, thus ordering these flexible loops nearby the catalytic site. The inter-switch motif, however, is retracted into the hydrophobic core. The ordered the switch motifs in GTP bound "ON" state thus form an conformation that are ready to interact with downstream effectors, while

the disordering in GDP bound "OFF" state disrupts the interactions (Wittinghofer and Pai 1991, Vetter and Wittinghofer 2001). ROC so far has been shown to have similar guanine nucleotide binding, GTPase activity and conformational change induced by GTP/GDP binding. However, it should be noted that ROC is also unique in that it is a G domain in complex protein and likely form one functional unit with its C terminal domain COR. Also, the GTPase dead mutations (T1343V, R1398L) in Ras G proteins cannot abolish the activity of ROC, while the Ras-like mutations (T1343G, R1398Q) has little effect on the GTPase activity neither (Biosa, Trancikova et al. 2013). Thus, albeit an active G protein, ROC might have different functional and structural mechanism from typical Ras-like G proteins.

The crystal structure of dimeric ROC has shed light on the mechanism how GTP/GDP binding would trigger the observed dimer-monomer interconversion in soltuion. The dimeric interface is mainly mediated by the switch and inter-switch motifs in a pretzel-like conformation, in which the inter-switch forms a beta hairpin and extends into the hydrophobic core of the counter protomer. The inter-switch motif is flanked by the conformational flexible switch I and switch II, thus it is a latch for dimeric interaction, which flips in or out by conformational change of the switch motifs as the hinge.

Since we only have the GDP bound dimeric structure, it is still not clear how the GTP binding results in the dissociation of the two protomers. However, the dimeric structure indeed provides clues on this dynamic process. The switch I and switch II "hinges" flanking the inter-switch "latch" are conformationally strained, in which S1360 and Q1411 have phi and psi angles of -93°, -90° and -105°, 96° respectively. Thus, the switch motifs are like spring-loaded "hinges" storing energy that may drive the conformational change

of the inter-switch "latch" when the dimer dissociates upon GTP binding. A potential process is that the γ-phosphate of the GTP interacts with the threonine and glycine residue in switch I and II respectively, which is the canonical mechanism in Ras-like G proteins. Based on sequence alignment and structural comparison by superimpose with human Ras (6Q21, GTP-bound), ROC indeed contains corresponding T1368 (T35 in Ras) in switch I and G1397 (G60 in Ras) in switch II (Figure 19C,19E and 23C). These conserved residues in both primary and tertiary structure indicate that their interaction with γ-phosphate of the GTP might trigger the switch motifs to off-load the energy and fold into the canonical Ras-like conformation, thus leading to the retraction of the flipped inter-switch "latch". Thus, it is consistent with the observation of dynamic dimerization of ROC in solution.

ROC is likely stabilized by the COR domain in the GTP-bound monomeric form. Due to the difficulty of obtaining enough protein of the full length LRRK, a lot of the interpretation on the potential functional mechanism of human LRRK2 is based on homology models, including ctROCO. The structural model (3DPU, 6HLU) of ctROCO is especially helpful to understand the potential interaction between the ROC and COR domains in human LRRK2. Thus, our next question is to test how the ROC structure of human LRRK2 is different from the homolog. Based on the model of ctROCO, it has shown an extensive interaction between the ctROC and ctCOR domains *in cis* and a constitutive dimeric interface mediated by COR (Gotthardt, Weyand et al. 2008, Gilsbach, Ho et al. 2012). It also implied a potential GAD mechanism of ctROC, which delivers an arginine finger into each other's catalytic site. However, this arginine residue is absent in human ROC, and ctROCO was recently reported to also dynamically convert between dimer and monomer in a similar mode as we observed in ROC (Deyaert, Wauters et al.

2017). Thus, the structural and functional mechanisms on ROC-COR tandem domains are still unclear.

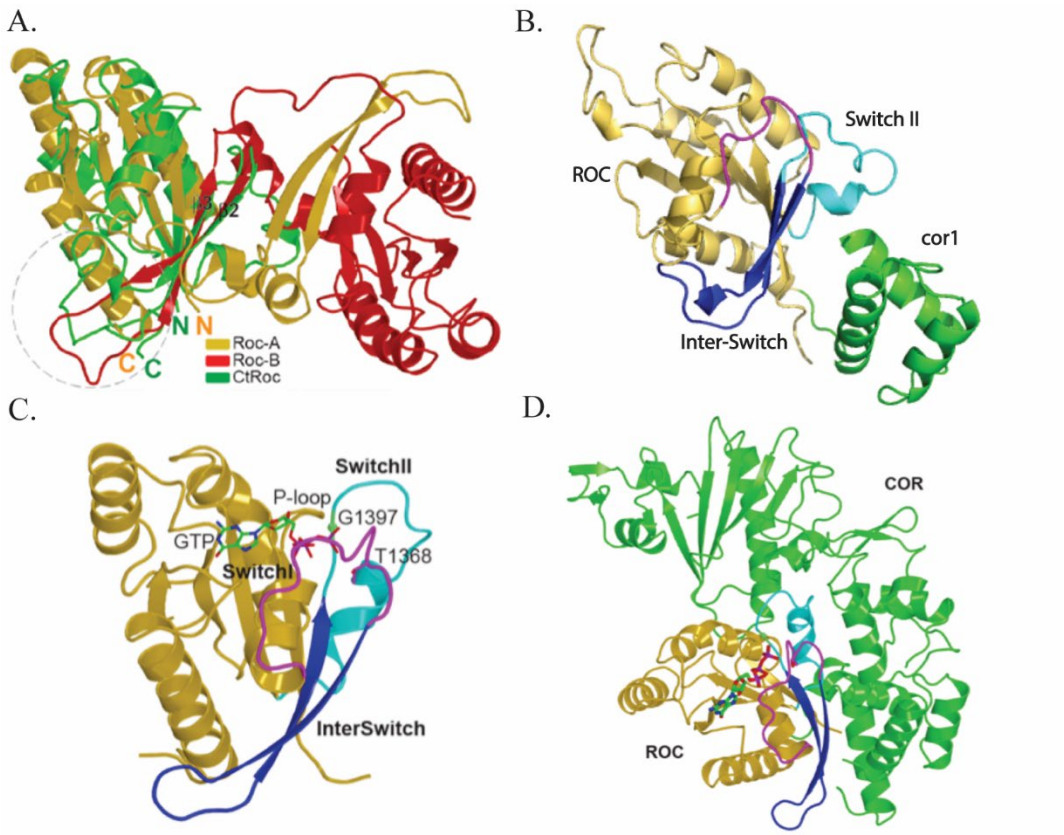


Figure 23. COR domain is likely stabilizing the monomeric state and thus regulating the dimerization dynamics
A). Superposition of human ROC dimer (gold and red) with the ROC domain of ctROCO (green) to compare the structures. The ctROCO has a Ras-like conformation, of which the overall folding is also similar to ROC. **B).** A homology model of ROC-cor1 (1329-1580) construct based on the crROCO structure (3DPU), which reveals the extended cor1 (1521-1580) regions is potentially interacting with switch II and inter-switch motifs of ROC. **C).** A monomeric homology model of ROC based on GppNHp-bound Ras (6Q21) D). Docking of the Ras-based homology model of ROC monomer to the COR domain of ctROCO (green), showing that the surface of switch II (cyan) and inter-switch (blue) is complementary to the COR domain of ctROCO.

To investigate the similarity and differences between human ROC and ctROCO, we first superimposed the two ROC domain structures (Figure 23A). Although the sequence identity is low (25 %), the overall conformations are similar (RMSD 2.2 $Å^2$). However, it should be noted that the switch I, inter-switch and switch II motifs are remarkably different that ctROCO adopts the canonical monomeric GTPase conformation, which retracts the inter-switch backward instead of flipping out. It should be noted that the inter-switch and switch II motifs are stabilized by the *in cis* interaction with its COR domain. Consistently, one of our extended human ROC constructs, ROC-cor1 (1329-1580), indeed has shown a prevalent monomeric conformation in solution, thus suggesting the COR domain is likely stabilizing ROC in the GTP binding induced monomeric conformation. Indeed, the homology model of the ROC-cor1 based on the ctROCO structure (3DPU) suggests that the cor1 extension may interact with switch II and inter-switch motifs (Figure 23B). To test this hypothesis, we also made a monomeric homology model of ROC based on GppNHp-bound Ras (6Q21) and docked it with the COR domain of ctROCO, which showed a potential complementary interface between the two domains. Thus, it is likely a common mechanism between human ROC and ctROCO, since both have similar dimer-monomer interconversion observed in solution, which consistently cycles between GDP-bound dimeric form and GTP-bound monomeric form.

Taken together, we hypothesized a potential structural mechanism of ROC-COR: the switch motifs and inter-switch mediate the dynamic dimerization (likely a local effect in ROC) through a “hinge” and “latch” mechanism. When bound with GDP, the inter-switch is flipped out by “spring-loaded” switch motifs and mediates the dimeric conformation; when bound with GTP, however, the γ-phosphate of GTP interacts with

residues in the “loaded hinges” of switch motifs and induces the flipping back of inter-switch “latch”, which is then fold into the Ras-like conformation with switch and inter-switch regions stabilized by its COR domain (Figure 23D). However, without a structure of the monomeric human ROC or a ROC-COR tandem bidomain, it still requires more work to confirm this potential mechanism.

C. PD-associated R1441 residue is critical for the dynamic interaction of ROC dimers

We have observed in solution that the disease-associated mutations R1441C, R1441G and R1441H results in a common reduction of GTPase activity and disruption of dimeric conformation. Thus, we hypothesized that the dynamic dimer-monomer interconversion is necessary for the GTPase function of ROC. Indeed, as suggested by the disulfide bond stabilized dimer, the dimeric conformation is stable in GDP-bound state but not stable in GTP-bound state, while "locked" dimer lose binding affinity to GTP, thus suggesting conformational change to monomer is likely required for GTP binding. However, it remains unclear how R1441C, R1441G and R1441H results the constant disruption of the dimeric conformation, and how absent of the dimeric state affects the GTPase functional cycles.

R1441 forms exquisite interaction with the switch II motif of the opposite protomer. To investigate the role of the R1441, we examined crystal structure and found that the side chain of R1441 extends across the dimer interface and interacts with the in loop between the distorted helixes (H2 and H3) in switch II region of the opposing protomer (Figure 18C and 19C). When we look closely at R1441 involved interaction, we find that the guanidinium group of R1441 forms π stacking with the phenyl ring of F1401, which orients the planar guanidinium group and locates its two ω-amines at ideal position to form hydrogen bonds with the backbone carbonyl oxygen of F1401 and the sidechain hydroxyl oxygen of T1404 (Figure 24A and 24B). The Van de Waals surface of the socket of this interaction is perfectly compensate for the R1441 side chain (Figure 24A). These

indicates that R1441 is part of an exquisite and strong interaction between the two protomer chains, thus showing a critical role in dimeric interaction.

Table 4. R1441 mutations and ROC oligomeric states

	Residue 1441	Hydrogen bond	π-stacking	Positive charge	Negative charge	hydrophobic
dimer & monomer	R, wild type	+		+		
	A					
dimer	Y	+	+			
	W	+	+			
monomer	K	+		+		
	H, pathogenic	+	+	+		
	N	+				
	C, pathogenic	+				
	S, pathogenic	+				
	T	+				
	G, pathogenic					
unstable	D	+			+	
	E	+			+	
	Q	+				
	I					+
unstable & insoluble	F		+			+
	V					+
	L					+
	M					+
	P					

Based on the stability and solubility of different R1441 mutations, the capability of forming hydrogen bond is likely essential. However, most mutations are not able to mimic the exquisite interaction of the wild type R1441, thus the hydrogen bond forming mutations R1441K/H/N/C/S/T only exist in monomeric form. The R1441Y/W are capable to form hydrogen bond and π-interaction through the aromatic side chain, thus mainly exist in dimeric form in solution. However, mutations of the opposite charge or large hydrophobic side chain make the ROC protein unstable in solution.

Given the exquisite and tight complementary interaction, it seems no other amino acid can replace the arginine. Indeed, we found the pathogenic mutations R1441C, R1441G and R1441H have only monomeric conformation in solution. To have a comprehensive understanding on the role of R1441, we substituted arginine to all different amino acids and tested their oligomeric state and stability. The results showed that mutations R1441K/H/N/C/S/G/T all lead to monomeric form in solution, while R1441Y/W mutants in contracts are mainly dimeric (Table 4 and Figure A.29). Albeit the secondary structure is not altered (Figure 24C), all R1441 mutations decrease the thermal stability of the ROC (Figure 24D), including the dimeric R1441W and R1441Y, thus indicating no amino residue can mimic the exquisite interaction by R1441. The dimeric R1441Y mutant was previously compared with the disulfide bond stabilized dimer, which reduces the GTPase activity, thus suggesting the monomeric conformation is required. Now with the structure model, the Y1441 may form a stronger π-stacking with F1401 as well as a hydrogen bond, however the different interaction may alter the conformation of the distorted helixes (H2 and H3), thus causing the reduction of thermal stability. In contrast, the monomeric mutations, including a likely conserved R1441K mutation (Figure 24E), reduce the GTPase activity (Figure 24F), thus suggesting the dimeric conformation is also necessary.

R1441 stabilizes the conformational strained switch II "hinge", thus modulating the dynamic switch motifs "hinges" and inter-switch latch in the dimer-monomer equilibrium. As discussed in Chapter 2.2B, the distorted helixes (H2 and H3) in switch II motif (ranging from G1397 to Q1411) forms the hydrophobic interaction with the opposing chain, of which the interaction is capped by hydrogen bonds of the two R1441 residues between the two chains in ROC dimer (Figure 18C and 19C). The distorted

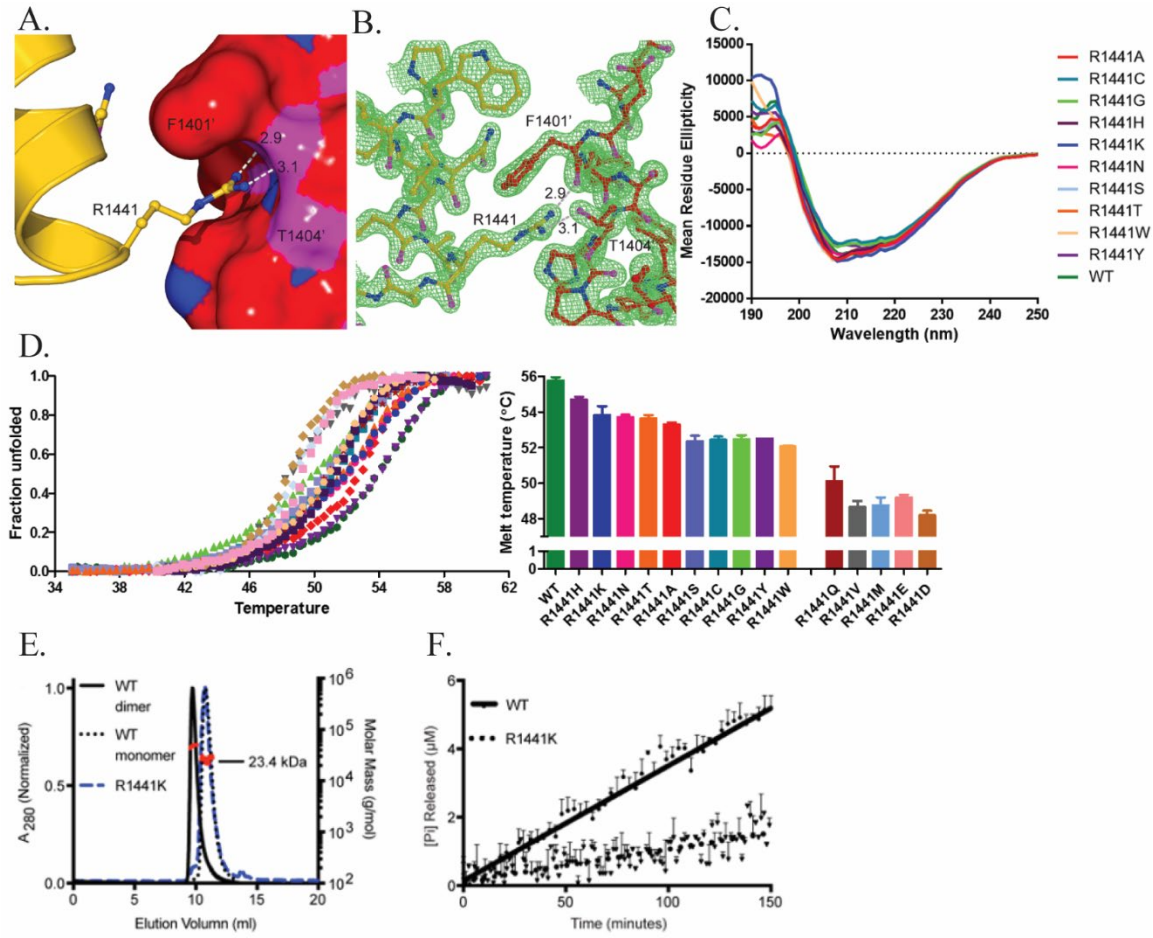


Figure 24. Comprehensive analysis of the function and conformation of the disease related residue R1441 **A).** Model showing the exquisite interaction mediated by R1441 to the F1401 and T1404 of the other protomer. **B).** Density map (σ=2.0) showing the R1441 mediated interaction. **C).** Circular Dichroism analysis of R1441 mutations shows the secondary structure is not altered. **D).** Thermal shifting assays on the R1441 mutations show reduced stability. **E).** SEC-MALS determined that the size of R1441K mutant in solution is 23.4 kDa, thus suggesting it is in the monomeric form. **F).** Side by side GTPase activity comparison showed the R1441K mutant has significantly reduced GTPase activity.

helixes are likely spring-loaded “hinges” with conformationally strained residue (Q1411 has phi and psi angle of -105° and 96°), thus R1441 mediated interaction is likely stabilizing the strained conformation of the switch II “hinge” in dimer. Together, these data suggest that R1441 mediates and may also modulate the dynamic dimer-monomer interconversion through interacting with the “spring-loaded” switch II “hinge”, thus helping stabilize and hold the loaded energy in dimeric form. This stabilizing effect may also maintain the structural plasticity and rigidity of the switch II region during the vast conformational change. Collectively, this potential role of R1441 in the “hinge’ and “latch” hypothesis is consistent with the observed dimeric disruption, reduced GTPase activity and thermal stability. Thus, R1441 is likely uniquely required to mediate the dynamic dimer-monomer equilibrium, which is a required process in the GTPase reaction cycles of ROC, while the disease-associated mutations disrupt the exquisite interaction by the guanidinium group of R1441, thus leading to a trapped GTP-bound monomeric form of ROC.

However, the structure of GTP-bound monomeric form of ROC is still unknown, thus suggesting the hypothesized mechanism of the GTPase reaction coupled cycles of dimer-monomer still need confirmation by future studies on ROC and COR domains. Also, it is essential to understand how the functional dynamics of ROC modulate the function of full length LRRK2. Thus, we decided to pursue our studies using the full length LRRK2 protein.

Chapter 3. Characterization of the kinase activity and guanine nucleotide binding of full-length LRRK2

3.1 Introduction

A. ROC domain of LRRK2 may regulate its kinase activity

LRRK2 mutations cause autosomal dominant PD, thus suggesting a gain-of-function effect. As discussed in previous chapters, the isolated ROC domain has dynamic dimerization coupled with GTPase function. The pathogenic mutations disrupt the GDP-bound dimeric conformation and reduce GTPase activity, thus leading to a prolonged monomeric GTP-bound form which enhances the downstream signaling. Thus, the biochemical and biophysical features observed from the recombinant ROC domain are consistent the gain-of-function effects of the pathogenic LRRK2 mutations in familial PD.

However, it is still unclear if GTP and GDP binding coupled dimerization of ROC also mediates the oligomeric states of the full length LRRK2. As a large complex protein, the other domains of LRRK2 may interact with each other and potentially contributes to the dimerization dynamics of LRRK2 as well. Thus, the biochemical and biophysical consequences of ROC mutations are still unclear on full length LRRK2. To have a comprehensive understanding, it is necessary to test our hypothesis on the full length LRRK2.

Kinase activity of LRRK2 is essential in the pathogenic mechanism of PD. As was discussed in Chapter 1.1C, the most common PD-associated mutations are found in the ROC-COR-Kinase tri-domain, thus suggesting both GTPase and kinase activities are essential for the physiological and pathological function of LRRK2. The most prevalent mutation G2019S is predicted to locate in the activation loop of the kinase domain, which

has been reported by multiple literatures to enhance the kinase activity of LRRK2 and reduce the viability of neuronal cells (West, Moore et al. 2005, Greggio, Jain et al. 2006). A crystal structure of LRRK2 homolog from *D. discoideum*, ddROCO4, reveals that residue G1179 (equivalent to G2019 in LRRK2) is part of the DFG (DYG in LRRK2) metal binding motif, while the mutated S1179 forms hydrogen bond with a nearby residue R1077 (equivalent to Q1919 in LRRK2), thus stabilizing the DFG loop to interact with a Mg^{2+} (Gilsbach, Ho et al. 2012). Thus, the G2019S mutations in LRRK2 is likely stabilizing the binding of Mg^{2+} in the catalytic site, thus enhancing the kinase activity and toxicity to neuronal cells.

However, the physiological substrate of the LRRK2 kinase is still unknown. The current assays for LRRK2 kinase are based on either autophosphorylation or peptide derived from general kinase substrate. The most commonly used artificial peptide substrate for LRRK2 is LRRKtide (RLGRDKYKTLRQIRQ), which is derived based on phosphorylated region of human ezrin (amino acids 561-573), moesin (amino acids 539-553) and radixin (amino acids 558-570) (West, Moore et al. 2007, Anand, Reichling et al. 2009). Thus, it should be cautious that the physiological kinase substrate of LRRK2 may require specific protein-protein interaction. Thus, kinase activity results based on LRRKtide may provide little information on the potential regulation of kinase substrate accessibility.

GTPase function may regulate the kinase activity of LRRK2. LRRK2 is dimeric in solution (Greggio, Zambrano et al. 2008, Klein, Rovelli et al. 2009), and the kinase activity is likely associated with the oligomeric states of LRRK2 (Sen, Webber et al. 2009, Berger, Smith et al. 2010). In cellular environment, LRRK2 adopts monomeric

state in cytosol but forms dimeric and higher oligomeric conformations at membrane surface (Berger, Smith et al. 2010, James, Digman et al. 2012). Thus, it suggests a potential regulatory mechanism on the kinase activity through dynamic dimerization of LRRK2. According to our observations on the isolated ROC domain, this dynamic dimerization equilibrium is likely mediated by GTP and GDP binding cycles.

Consistently, it was reported that R1441G, R1441C mutations in the ROC domain cause both reduced GTPase and enhanced kinase activity (Guo, Gandhi et al. 2007, West, Moore et al. 2007), thus suggesting that the reduction of GTPase activity by mutations in ROC also cause up-regulated kinase activity. Further studies have shown that mutations K1347A, T1348N in the nucleotide binding P-loop disrupt guanine nucleotide binding, while also dramatically decrease the kinase activity (Ito, Okai et al. 2007, West, Moore et al. 2007, Jean-Marc Taymans 2011, Biosa, Trancikova et al. 2013), thus suggesting the kinase of LRRK2 is dependent on guanine nucleotide binding. Collectively, these data all suggest that ROC may regulates the kinase activity of LRRK2.

B. Rationale and overview of this chapter

Using a stable recombinant construct of the ROC domain, we have found a unique mechanism of dimer-monomer inter-conversion, which is coupled with the GTPase reaction cycles. Our crystal structure of a GDP-bound homodimer of ROC suggests that the the dimeric interfaces are mediated by switch I, switch II and inter-switch motifs, thus conformational changes induced by the exchange of guanine nucleotides would trigger the dynamic dimer-monomer interconversion. The disease-associated mutations R1441C, R1441G and R1441H disrupt the dimeric conformation and reduces the GTPase activity. Thus, the pathogenic mutations trap the ROC domain in a prolonged GTP-bound monomeric state, which causes an extended "ON" signal for the downstream effectors. However, it is still not clear if the same mechanism also happens in full-length LRRK2, in which ROC might potentially regulate its own kinase activity as an "internal molecular switch". Thus, we decided to develop a capable system to express and purify full-length LRRK2 and test the biochemical function of both kinase and GTPase.

This chapter describes our overall goal to establish a robust expression and purification system of LRRK2 full-length protein, and examine the biochemical function of the kinase and GTPase domains. To obtain a LRRK2 protein sample with sufficient purity and quantity, we utilized the BacMam virus system, and expressed the full-length LRRK2 in adhesion culture of HEK293T cells. To further improve the efficiency, we also tested the BacMam system in Expi293 cell suspension culture. Comparing the adhesion and suspension culture systems, the suspension culture yields higher amount of protein while decreasing the consumption of plastic flasks and labor time. Both purified LRRK2 from adhesion HEK293T and suspension Expi293 cells showed good purity for kinase

assays. However, the protein yield is still too low to perform GTPase activity assays or crystallization condition screens for structure determination.

Using LRRK2 samples from BacMam expression system, we tested the kinase activity of LRRK2 using a general peptide substrate LRRKtide. The purified LRRK2 protein sample showed robust active kinase activty. The specific activity is 0.193 min^{-1} based on a linear titration of LRRK2 protein sample (ranging from 0.007 μM to 0.111 μM) against a kinase reaction system of 0.25 mM ATP and 0.1 mM LRRKtide, thus suggesting an average of 0.193 γ-phosphate is transferred onto LRRKtide by each enzyme per minute.

To investigate if the ROC domain regulates the kinase activity through GDP/GTP cycling, we compared the kinase activty of LRRK2 incubated with GDP, GTP, or non-hydrolysable GTP analog GppNHp. However, we found no significant difference in the kinase activity between the samples. To further investigate if this is concentration dependent, we tested the kinase activity again by pre-incubating LRRK2 with different titers of GDP or GTP ranging from 0 to 1.0 mM. However, the kinase activity still showed no significant difference between GDP or GTP incubated LRRK2. Thus, our result suggests that the kinase activity of LRRK2 is likely independent to the exchange of GDP/GTP in the ROC domain. However, the guanine nucleotide binding is likely still required for the kinase domain function. Indeed, mutations in P-loop, which disrupt guanine nucleotide binding, were reported to greatly reduce kinase activity (Liu, Dobson et al. 2008, Jean-Marc Taymans 2011, Biosa, Trancikova et al. 2013). Thus, the kinase activity of LRRK2 is likely independent to GDP/GTP exchange but dependent on the capability of guanine nucleotide binding at ROC.

3.2 Materials and Methods

A. Expression of LRRK2 in HEK293T cells

To express the full length LRRK2 in HEK293T cells, we utilized the BacMam virus expression system. BacMam virus is modified insect cell virus (baculovirus) that it contains recombinant mammalian expression cassette for transgene expression in mammalian cells. The BacMam virus particles can be taken up by mammalian cells through endocytosis. The double strand DNA of the BacMam is then released in the cytosol and migrate to the nucleus for transcription and expression. Thus, the BacMam expression system can serve as a vehicle to efficiently transfect genes into the mammalian cells for expression. The BacMam virus is safe to use because it cannot replicate itself in mammalian cells. Also, the virus can easily be prepared using insect cell cultures.

For the expression of the LRRK2, the original BacMam virus (P1) containing 3 × flag tagged full length LRRK2 gene was obtained from Life Technologies. To express the target LRRK2 protein, the host HEK293T cells were scaled up by growing to confluent in 4 T175 flasks in the DMEM high glucose medium. The final cell count in the resuspended cell solution was around 140 million cells total. The cells solution was diluted with the culture medium to density of 3.5×10^6 cells/ml. To transfect the HEK293T cells with the BacMam, 0.175 ml of the gross virus solution per million of the cells counts was added into the suspended cells with addition of BacMam Enhancer solution (Thermo Fisher), and then incubated together for 25min in room temperature. The Enhancer solution enhances the expression level of LRRK2 (Figure A.30A). Then, additional fresh DMEM was added in to make the cells density for plating as 0.56×10^6 cells/ml. Then the cells were plated into total of 1400 cm^2 (8 T175 flasks) adhesion culture, in which the plating density is around

0.1×10^6 cells/cm^2. The cells were then continue cultured for 48hrs. The cells were then harvested by treating with Versene solution (Gibco, containing no protease) 2min in 37˚C to release the adhered cells from the flask. The cells were then centrifuged with 300×g to harvest and then stored in -80˚C.

B. Preparation of BacMam virus solution in Sf9 insect cells

The baculovirus solution of BacMam is prepared in insect cell line Sf9. The cells were cultured in ESF921 (Expression Systems) media. The culture was passaged with cycles of 3 days in suspension culture. The cells were plated with density of 0.5×10^6 cells/ml on day 1 and passaged on day 3, when density reach to around 5×10^6 cells/ml.

For generating the BacMam virus passage 2 (P2) for LRRK2 expression, the P1 virus was added 1:10 ratio into the suspended Sf9 cells in ESF921 media, which were plated with cells density of 0.4×10^6 cells/ml. After continue culturing for 48hrs, the culture solution was centrifuged at 1000 × g, 10 min to separate the insect cells (pellet) and the gross virus solution (supernatant). The expected Sf9 cells viability should be more than 85 % so that the gross virus solution would not be contaminated by debris of dead insect cells.

C. Adaptation of BacMam expression system into Expi293 suspension culture

To express LRRK2 in shaking culture system, we adopted the Expi293 expression system (Thermo Fisher), which was developed based on HEK293 cells and designed for efficient bulk protein expression. The Expi293 cells can be revived directly from frozen stock into suspension cell culture using Expi293 expression medium (Thermo Fisher), however it would have low initial viability and cost additional passages to recover. Alternatively, reviving the cells in adhesion culture using DMEM medium with FBS would

generate a robust passage of Expi293 cells with high viability (Figure A.30B). After one passage in adhesion culture, the cells were then transferred into the FBS free Expi293 expression medium. Expi293 cells were plated with cell density of 0.5×10^6 cells/ml, and the density would reach to 2.5×10^6 cells/ml after 3 days culture in 37℃, 8 % CO_2 and 128 rpm (Figure A.30C).

To express LRRK2, the cells were harvested and re-suspended in new media with concentrated density of 3.5×10^6 cells/ml (volume of 40 ml in typical scale) mixing 2:1 with BacMam virus solution (volume of 20 ml in typical scale). The total of 60ml of concentrated cell-virus mix are incubated in room temperature for 30 min without shaking. After the incubation, another 120 ml of expression media are added to make the final volume of 200 ml. Thus, the transfected cells are plated with density of 0.7×10^6 cells/ml and 10 % volume of virus solution. The transfected cells were then cultured another 2 days for protein expression. Likely due to the toxicity of LRRK2, the proliferation of cells was paused after the transfection and the observed density upon harvesting is around 0.6×10^6 cells/ml.

D. Protein purification of LRRK2

To purify the LRRK2 protein from either HEK293T or Expi293 cells, the cells pellet was resuspended in buffer containing 50 mM Tris-HCl pH7.4, 150 mM NaCl, 0.5 mM EDTA, 5 mM $MgCl_2$, 0.1 % BOG, 2 mM DTT, 10 % Glycerol with 1 × HALT protease inhibitor cocktail (Thermo Scientific). The cells were then lysed by homogenization. The insoluble fraction was separated from the soluble lysate by ultracentrifugation 35000 rpm, 30 min. Because the LRRK2 protein was cloned with a FLAG N-terminal tag, the soluble lysate was then incubated with anti-FLAG resin (ANTI-

FLAG M2 affinity gel, Sigma-Aldrich) for 1 hour in 4˚C. After the incubation, the resin was transferred to an open column and let the supernatant flow though. The protein was then eluted off from the resin with 0.15 mg/ml 3 × FLAG peptides solution. The protein eluate was then polished with anion exchange column (HiTrap Q FF, GE) to remove the flag peptides in the solution. The final protein solution was concentrated and frozen in liquid nitrogen, and then stored in -80˚C.

E. Western blot

To estimate the expression level of the LRRK2 protein in the HEK293T and Expi293 cells, BacMam transfection was performed on the cells in small culture scale: HEK293T in T75 flask and Expi293 in 50 ml suspension culture. The culture and transfections were the same as mentioned in the above A section. After the cells were harvested, the cells were resuspended in buffer containing 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.5 mM EDTA, 5 mM $MgCl_2$, 0.1 % BOG, 2 mM DTT, 10 % Glycerol and sonicated 15 sec. The lysate was then centrifuged 16000 × g, 10 min to separate the soluble and insoluble lysate.

The samples were then run on SDS PAGE and transferred to PVDF membrane for western blot. The rabbit anti-ROC domain antibody was used as the primary antibody to detect the LRRK2 protein band. The IRDye 800CW goat anti-rabbit IgG (Li-COR) was used as the secondary antibody. The western blot membrane was scanned on the Odyssey Imaging system (Li-COR).

F. Kinase activity assay

A radionucleotide based assay was used to measure the kinase activity assay. The [γ-32P] ATP was used to track the transfer of the γ phosphate from the ATP to the substrate:

LRRKtide (RLGRDKYKTLRQIRQ, Signalchem), which is derived from human ezrin (amino acids 561-573), moesin (amino acids 539-553) and radixin (amino acids 558-570).

12 ng/µl of the LRRK2 protein was mixed with 0.2 mg/ml LRRKtide and 0.25 mM ATP (include 0.25 mCi/ml of the [γ-32P] ATP) in buffer containing 5 mM MOPS, pH 7.2, 2.5 mM B-glycerol-phosphate, 4mM $MgCl_2$, 2.5 mM MnCl2, 1 mM EGTA, 0.4 mM EDTA, 0.05 mM DTT. A total reaction system of 25 µl was prepared and incubated at 30˚C 15 min for the kinase reaction. The reaction solution was then transferred to a P81 paper and dried. The dried P81 paper was then rinsed tandemly three time in 1 % (v/v) phosphoric acid solution. The P81 paper was then put on the vacuum manifold to remove the remaining phosphoric acid solution. After dried, the P81 paper was transfer into a vial with 5 ml of scintillation cocktail (PerkinElmer). The radioactivity was then counted on the Beckman LS6500 scintillation counter. The total phosphorylation of the substrate was then calculated based on the standard curve of the [γ-32P] ATP.

3.3 Results and Discussion

A. Robust expression of full-length LRRK2 using a BacMam expression system

The bottleneck of the study of LRRK2 is in obtaining samples with enough quantity and purity. As a complex protein of 2527 amino acids, LRRK2 has multiple individually folded domains and its tertiary structure is expected to be more complex than other small single domain proteins. Given its large molecular size, it is not surprising that we had found that the full-length protein cannot be recombinantly expressed in *E. coli*. Also, due to its kinase activity, overexpressing of this protein is often toxic to the host cells.

In order to obtain enough LRRK2 for our studies, we utilized the BacMam expression system to overexpress the LRRK2 protein in HEK293T cells. BacMam is a modified baculovirus expression system that contains a mammalian gene expression cassette inserted in the baculovirus genome. The baculovirus is not able to "actively" infect the mammalian cells, however, they can be taken into the cells via endocytosis of the host cell (Figure 25A, 25B). After internalized into the cell, the recombinant target gene will be relocated to the nucleus by a yet unknown mechanism, and the target protein product will be made by the host cell's protein expression system.

LRRK2 can be robustly expressed in HEK293T cells. Using the BacMam for the delivery of the LRRK2 gene, we used HEK293T cells in the traditional adhesion culture. The proliferation of the HEK293T cells was arrested after the transfection of LRRK2 by BacMam virus, which is likely due to the toxicity of the overexpressed LRRK2. However, we didn't observe any significant cell death, thus suggesting that the HEK293T cells were available host for the overexpression of LRRK2. The cells were harvested 48hrs after transfection and then the LRRK2 protein was purified using anti-flag affinity purification.

To test the purity, the LRRK2 samples were ran on SDS PAGE. A clear single band above the 250 kDa protein marker can be observed in the lane of the LRRK2 purified sample, whose identity was confirmed with anti-LRRK2 antibody by Western blot (Figure 25D),

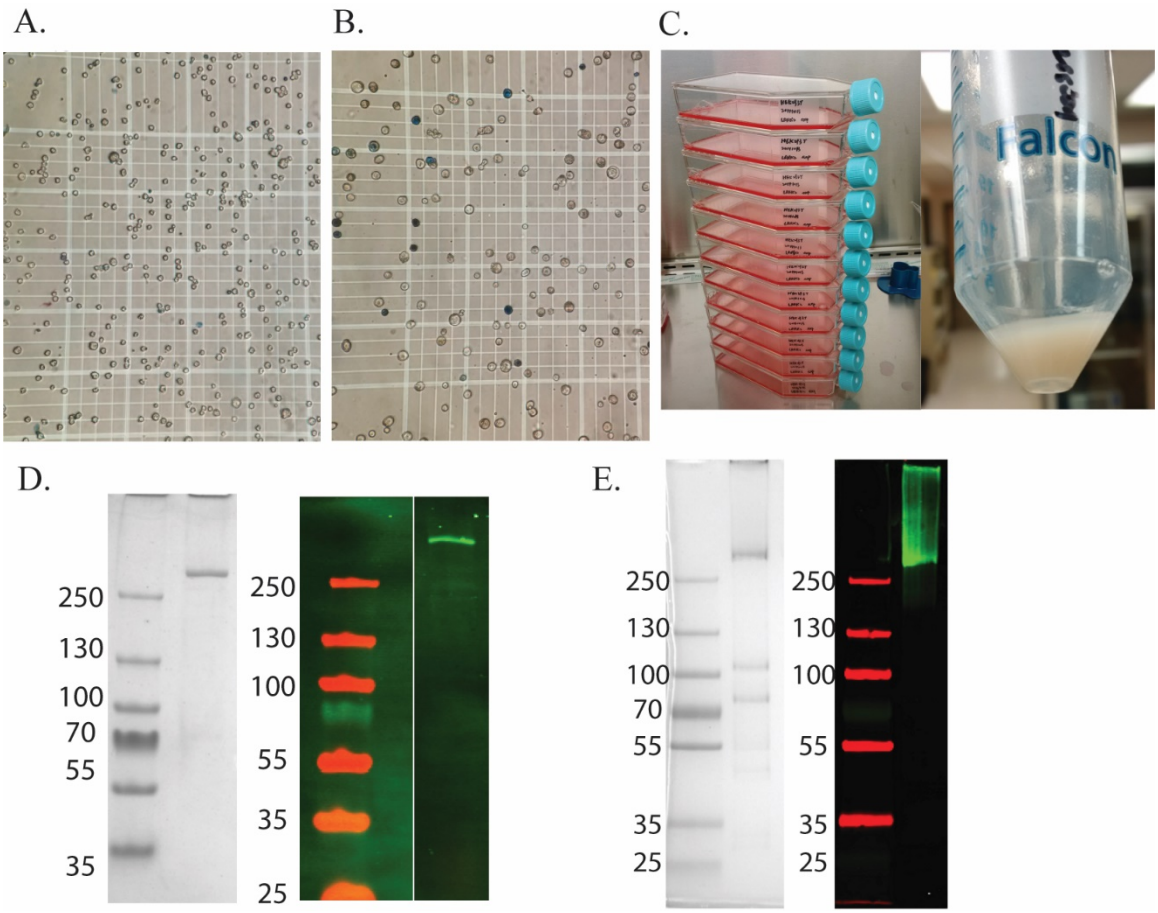


Figure 25. Expression of LRRK2 in HEK293T and Expi293 cells using BacMam expression system **A).** Cells before the BacMam transfection. **B).** Cells after transfected BacMam for 2 days. The cells showed swelled size after BacMam transfection, which is similar to insect cell baculoviral expression system. **C).** An example of a batch of expression in the BacMam transfected HEK293T adhesion cells. Total of 400 ml expression media was plated into 11 T175 adhesion culture flasks (total of 1925 cm^2). The plating density of cells is 0.1×10^6 cells/cm^2 (0.5×10^6 cells/ml). Total of about 2 ml cells paste was obtained. **D).** SDS PAGE (left) and western blot (right) of the purified full length LRRK2 protein sample from adhesion cultured HEK293T cells. **E).** SDS PAGE (left) and western blot (right) of the purified full length LRRK2 protein sample from suspension cultured Expi293 cells.

thus indicating LRRK2 was successfully expressed in HEK293T and purified. In a typical purification from 12 T175 flask totaling 420 ml media, about 50 µg of LRRK2 protein were obtained from about 2 ml cells paste (Figure 25C). However, to purify large amount of LRRK2 using the HEK293T cells would require large adhesion culture limited by the surface area of the culture system, which made it laborious and expensive. To improve the expression efficiency and reduce cost, we tried to adapt the expression system into shaking suspension culture, which allows for easier large-scale protein production.

Adaptation of BacMam system into suspension culture. Expi293 (Thermo Fisher) is a novel cell line designed for bulk protein expression in suspension culture, which is developed based on HEK293 cells. Compared to the conventional adhesion culture, the suspension culture benefit from higher yield of cells from the same amount of expression media. Also, instead of consuming many surface treated adhesion culture flasks, the shaking culture only requires re-usable Erlenmeyer flasks.

The BacMam virus transfection was designed similarly to the method for HEK293T cells. For transfection, the virus solution was mixed (1:2) with high density cells in expression media (3.5×10^6 cells/ml) and incubated together without shaking for 30 min. After the transfection, the cells were then plated with additional media so that the cells are plated with suitable density (0.7×10^6 cells/ml) for expression in suspension culture. The cells were then continue cultured for 48 hours for protein expression. Similar as it was observed in HEK293T cells, the proliferation of Expi293 cells was halted after the transfection. In a typical scale of 200 ml culture of transfected cells, total harvested cell paste is around 2 mls. Thus, the cells paste amount (per unit of media volume) generated

from the Expi293 suspension culture is about twice as much as from HEK293T adhesion culture.

The LRRK2 protein was purified from Expi293 cells using the same method. A protein band above the 250kDa marker were shown in SDS PAGE (Figure 25E), thus suggesting the LRRK2 protein were successfully expressed in Expi293 cells. The final yield of LRRK2 protein was around 60 µg from the 200 ml culture. Comparing the two expression systems, it seems the Expi293 suspension culture has higher yield of protein than the conventional HEK293T adhesion culture. The purified LRRK2 protein amount was around 30 µg per ml of cells paste in Expi293 cells from shaking culture, while it was about 25 µg per ml of cells past of HEK293T cells from adhesion culture. Given the variation of yield between batches, the LRRK2 expression level per unit of cells are likely the same in the two expression systems. However, there are about twice as much of cells obtained from the shaking culture, therefore the protein yield is also about double in the Expi293 shaking culture per volume of expression media. Thus, the Expi293 expression system is an efficient and robust way to express LRRK2 full-length protein in bulk.

B. LRRK2 kinase activity and potential effects of guanine nucleotide binding

Up-regulated LRRK2 kinase activity is a common cause of PD. Previous studies in cells and immunopurified LRRK2 samples have shown that this kinase domain mutation causes up-regulated kinase activity. However, due to the limited amount and purity of the protein, the kinetics detail of the kinase activity of LRRK2 remains not well understood.

The LRRK2 expressed and purified from HEK293T has active kinase activity. We have obtained LRRK2 full-length protein sample with good purity from the BacMam expression system in HEK293T cells. To test if the LRRK2 protein purified from BacMam-HEK293T expression system is active, a radioactive based kinase assay was performed. The phosphorylation was tracked by P32 labeling on the γ-phosphate of ATP. Since the physiological substrate of LRRK2 is still unknown, a small peptide (RLGRDKYKTLRQIRQ) called LRRKtide was used as the substrate in the assay, which was derived from general kinase substrates ezrin/radixin/moesin (ERM). The LRRKtide has been reported to be a potent general substrate for LRRK2 (West, Moore et al. 2007, Anand, Reichling et al. 2009).

To determine the specific activity of LRRK2, kinase assay is titrated with different concentration of LRRK2. Our result showed a clear linear correlation between the total phosphate transferred to LRRKtide to the titer of the concentration of the LRRK2 sample, thus indicating that the purified LRRK2 is an active kinase that has a specific activity of 0.193 min^{-1} (Figure 26A). The specific activity is defined as the amount of γ-phosphate of ATP transferred onto LRRKtide per minute by each molecule of the LRRK2 enzyme.

The kinase activity of LRRK2 is dependent on guanine nucleotide binding, but not the exchange of GDP/GTP. Our next question is to investigate whether the GTPase

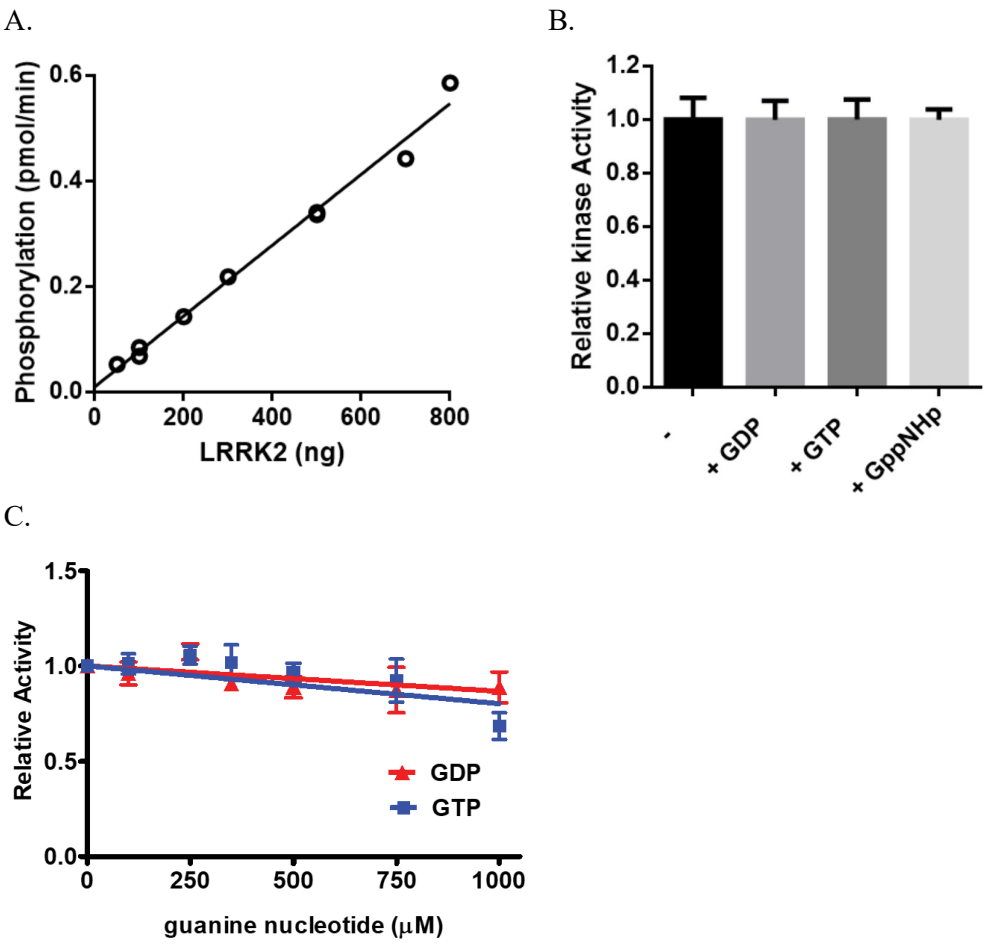


Figure 26. Kinase activity of LRRK2 and effects of guanine nucleotides
A). Phosphorylation (pmol/min) of the kinase plotted against LRRK2 amount (ng) in the 25 µl reaction. The slope of the 0.00067 pmol/(min·ng) represents the specific kinase activity of LRRK2, calculated as $0.193 min^{-1}$. Which means average of 0.193 reactions happened on each LRRK2 molecule per minute. Thus, the purified LRRK2 protein from BacMam expression system is active kinase. **B).** Comparison of kinase activity of LRRK2 after incubation of 1mM GDP, GTP or GppNHp. However, no significant difference on kinase activity was observed. **C).** Comparison of kinase activity of LRRK2 incubated with GDP/GTP ranging from 0 to 1 mM. However, still no signification difference was observed between GDP or GTP incubated LRRK2 kinase activity. With increased guanine nucleotide concentration, the kinase activity was found to slowly decrease, which is likely due to the unspecific competition of guanine nucleotide to the ATP (250 µM) in the reaction system.

function of ROC affects the kinase activity of LRRK2. As discussed in previous Chapters, we have observed a dynamic dimer-monomer equilibrium of ROC coupled with the GTPase reaction cycles. Thus, we were hypothesizing the vast conformational change between GDP and GTP bound forms of ROC might regulate the function of its own kinase domain. To test this, we preincubate the LRRK2 protein with either GTP, GDP or GppNHp (non-hydrolysable analogue of GTP) respectively, and then compare their kinase activity. However, the results showed no difference in activity between the LRRK2 preincubated with GTP, GDP or GppNHp (Figure 26B).

To test if this is due to insufficient dosage, LRRK2 was then titrated and incubated with different amount of GDP or GTP ranging from 0 to 1.0 mM (Figure 26C). The result shows no significant difference in kinase activity between GDP and GTP titrated LRRK2. The kinase activity shows a small reduction in the high concentration of guanine nucleotide, which is likely due to an unspecific competition of GDP or GTP to the substrate ATP (0.25mM). Thus, our data suggests that the kinase activity of LRRK2 is likely not directly depending on the guanine nucleotide cycles of the GTPase domain.

Comparing with literatures, although there are reports on GTP and GDP binding regulation on kinase activity (Guo, Gandhi et al. 2007, West, Moore et al. 2007), but it remains controversial that other groups have shown, similar to our observation, that GDP, GTP or non-hydrolysable GTP analogues GTPγS and GppCp binding has no alteration on the kinase activity of LRRK2 (Liu, Dobson et al. 2008, Jean-Marc Taymans 2011, Biosa, Trancikova et al. 2013). However, these reports have consistently shown that mutations in the P-loop of ROC (K1347A, T1348N), which disrupt the capability of guanine nucleotide binding, completely kill the kinase activity of LRRK2. Thus, LRRK2 has shown a guanine

nucleotide binding dependent kinase activity, but it seems only require the capability of guanine nucleotide binding while is independent to the GDP/GTP binding cycles.

However, the potential caveat on this conclusion is that, the physiological substrate for LRRK2 kinase is still unknown. Thus, the general peptide substrate LRRKtide cannot show the substrate specific effects. A potential factor is that the alteration on kinase activity is likely mediated by the conformational change of the substrate binding motifs, thus the effect can only be observed on the specific protein substrate. Also, protein interactors in the cells may interact with LRRK2 and modulate the regulation of kinase activity, thus the isolated LRRK2 protein sample lost the regulatory mechanism due to the absence of the interactor. This could be a potential explanation on why some studies (less pure LRRK2 but interactor containing) observed the regulation of ROC on kinase through GTP/GDP cycles, while our result and other groups did not observe the effects.

Lacking enough protein to perform more comprehensive analysis on full length LRRK2, we are not able at this moment to further investigate its biochemical and biophysical function. The future plan is to further optimize the expression system of LRRK2 in HEK293 cells, while also utilize insect cells expression systems to produce the protein more efficiently. We will investigate the oligomeric dynamics of LRRK2, its GTPase activity, and the kinase activity on different protein substrates. Also, with better quality and yield of full length protein, it will allow us to screen and determine the crystal structure of LRRK2, which will ultimately solve the current mystery of the functional mechanism of LRRK2, thus paving the way for understanding the pathogenesis of LRRK2-associated PD and therapeutic approaches of LRRK2-targeted drug design.

Chapter 4. Conclusion

Collectively, we have found an intriguing dimer-monomer inter-conversion coupled with GDP/GTP binding using a stable construct of the ROC domain: the GTP-bound ROC converts from dimeric to monomeric form; the GDP-bound converts from monomeric to dimeric form. Such inter-conversion is likely required for the biochemical function, since GTPase activity is reduced by mutations that disrupt the dimeric conformation. Also, a disulfide bond "locked" dimer showed reduced GTP binding affinity, thus suggesting the monomeric form is required for the binding of GTP.

To understand the structural mechanism, we have solved a 1.6 Å crystal structure of a GDP-bound homodimer of ROC, which suggests that the switch I, switch II and inter-switch motifs are directly responsible for the dimeric interaction. Different from the structure of Ras-like G proteins, the inter-switch motifs are flipped-out and exchanged in the dimer, which work like dual "latches" that engage and fasten the two protomers. The inter-switch "latch" is flanked by conformational strained switch I and switch II motifs, which thus act as spring-loaded "hinges".

Here we propose a model of the dynamic dimerization of ROC coupled with GTPase cycles. We have observed 3 different forms of ROC in solution: 1) a GDP-bound homodimer fastened by "latches" with strained "hinges"; 2) GTP-bound monomers with stabilized "hinges" and retracted "latch"; 3) GDP-bound monomers with flexible "latch" and "hinges", which is capable to form dimeric complex again. Upon the binding of GTP in the catalytic site, the γ-phosphate interacts with the switch I and switch II "hinges" and triggers a vast conformational change of the inter-switch "latch" to retract into the core of its protomer, thus forming a catalytically active monomer with Ras-like conformation,

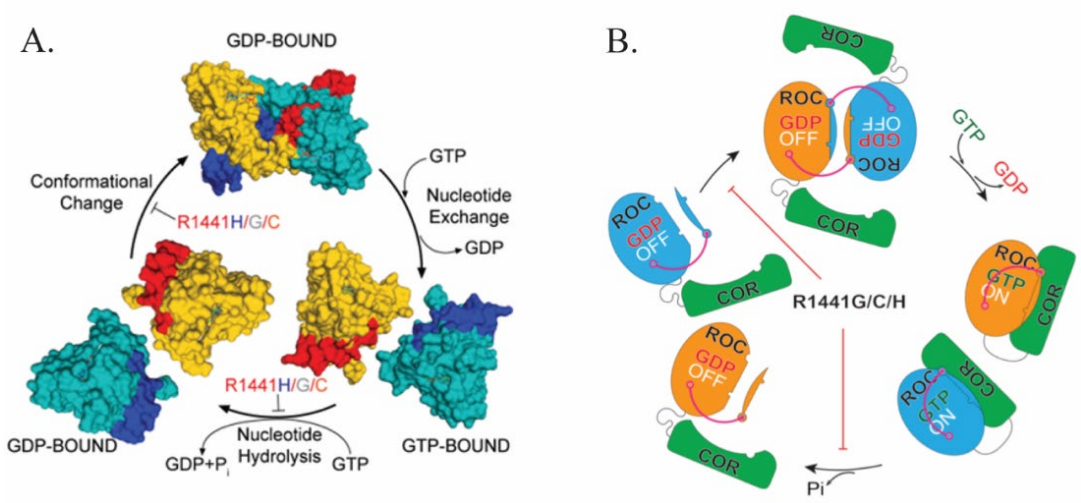


Figure 27. Model of the GTPase cycles coupled with dynamic dimer-monomer interconversion of the ROC domain of LRRK2

A). Model of the dynamic dimerization of ROC domain based on observation in crystal and solution. The dimer-monomer interconversion is regulated by the GTPase reaction cycles, of which the process is mediated by the conformational change of the inter-switch "latch" and switch motifs "hinges". GDP-bound ROC forms stable homodimer, which converts into monomeric form after GTP binding. In the process, the GTP interact and trigger conformational change of the loaded switch "hinges", thus flipping the inter-switch "latch" backwards to the monomeric core and forming a stable monomeric GTP-bound ROC, which is likely the active conformation that trigger the downstream signaling effect. The GTP is hydrolyzed to GDP, and likely form a transitional monomeric state of GDP-bound monomer, which then converts into GDP-bound dimeric conformation again. The pathogenic R1441C/G/H mutations cause reduced GTP hydrolysis and disrupt the dimeric conformation, thus trapping the ROC in the GTP-bound monomeric form and constant "on" state of the signaling. **B).** Potential model of ROC-COR functional unit. In the monomeric conformation of ROC, the COR domain may stabilize the retracted inter-switch "latch" and switch II "hinge" motifs. However, after GTP is hydrolyzed to GDP, the switch motifs become flexible and no longer interact with COR. The flexible inter-switch "latch" and switch motifs "hinges" then interacts between each other and form into the homodimer again.

which then initiate the hydrolysis of GTP. After GTP is hydrolyzed to GDP, ROC enters a GDP-bound monomeric state. Due to the absent of a γ phosphate in the catalytic site, the switch I and switch II "hinges" are likely highly flexible in solution. When ROC protomers are in a conformation with extended "latch", they then engage and form dimeric conformation again (Figure 27A). COR may interact and stabilize with inter-switch and switch II motifs of ROC in the GTP-bound monomeric state, while in contrast the interactions are excluded in the GDP-bound dimeric conformation of ROC. Thus, the COR

domain may involve in the GTPase reaction cycles and mediate functional consequence during the GTP-bound "ON" state of ROC (Figure 27B).

To understand if the GDP/GTP cycles coupled dimerization equilibrium is involved in PD, we have found that pathogenic mutations indeed disrupt the dimeric conformation of ROC in solution. As revealed by the dimer structure, the arginine side chain of R1441, the pathogenic related residue, stabilizes the latch motifs by forming two hydrogen bonds and π stacking across the dimeric interface near the switch II motif. Thus, the pathogenic mutations destabilize the latch motif, and consequently disrupt the dimeric interaction. Indeed, the R1441C, R1441G, R1441H mutations (as well as recently reported pathogenic mutation R1441S) consensually reduce GTPase activity and form only monomers in solution, thus resulting in a prolonged GTP-bound monomeric state of ROC, which is likely to generate an extended "ON" signal to the downstream cellular pathways of LRRK2 in PD.

ROC domain in the full-length LRRK2 may mediate similar oligomeric dynamics. Indeed, the LRRK2 homolog, ctROCO, has been shown to have the same dynamic dimer-monomer interconversion regulated by GDP/GTP binding (Deyaert, Wauters et al. 2017). However, such interconversion has not been observed in full-length human LRRK2. Due to the relatively small dimeric interface mediated by the latch motifs (switch I, inter-switch, switch II), it is not likely that ROC alone is enough to mediate the dimerization of the full-length LRRK2. Other domains, including COR and WD40, are likely more essentially involved in the dimeric interaction of LRRK2 (Gotthardt, K. et al. 2008; Zhang, P. et al. 2019).

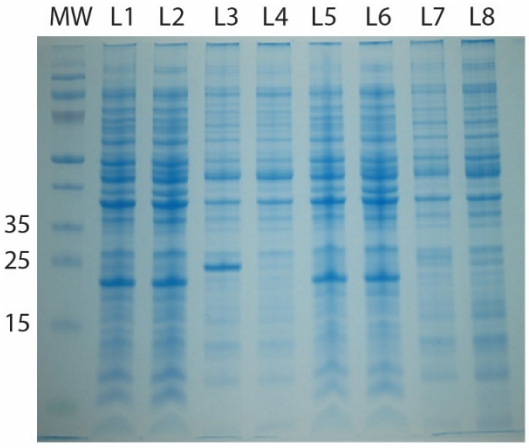
It is potentially possible that the ROC domain mediates a local conformational change rather than a vast oligomeric state conversion of the full-length LRRK2. In the GTPase reaction cycles, the GTP-bound monomeric ROC enables interaction with COR and triggers downstream signaling; while the GDP-bound ROC forms a fastened homodimer in its local region, thus excludes the interaction with COR (Figure 27B).

To further investigate the mechanism of full-length LRRK2, we have established a robust expression and purification system in HEK293T cells. The purified LRRK2 showed active kinase activity in solution. However, we did not observe any difference of kinase activity between GTP or GDP incubated LRRK2 in our initial test. Comparing with literatures, studies have shown similar observation, but mutations (K1347A, T1348N) disrupting GTP/GDP binding impair kinase activity (Liu, Dobson et al. 2008, Jean-Marc Taymans 2011, Biosa, Trancikova et al. 2013). Thus, the kinase activity of LRRK2 is likely not directly regulated by GDP/GTP exchange, but dependent on the capability of guanine nucleotide binding in ROC domain.

To sum up, our study has established a comprehensive understanding on the biochemical and biophysical features of ROC, the GTPase domain of LRRK2, as well as the structural basis for its unique dynamic dimerization. The observations in solution revealed that the GTPase activity is dependent on the dimer-monomer interconversion, which is disrupted by the pathogenic mutations R1441C, R1441G and R1441H, thus suggesting a novel strategy to regulate the GTPase of LRRK2 by targeting at the dimeric interaction. The 1.6 Å crystal structure of the dimeric ROC reveals a unique switch II and inter-switch motifs mediated dimeric interaction, thus paving the way for structure-based drug design and future therapeutic development for PD.

Appendix

Figure A.28. SDS PAGE of the expression tests of ROC constructs

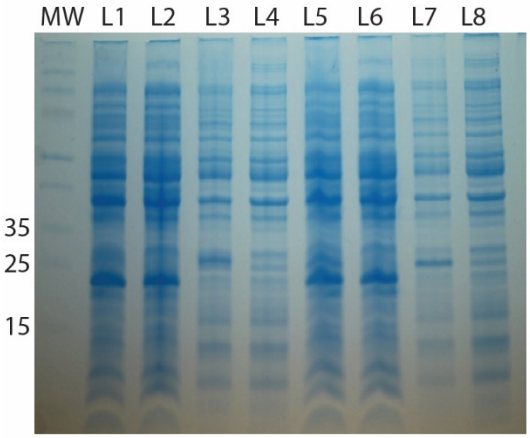


his-ROC 1329-1529:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

his-ROC 1315-1520:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

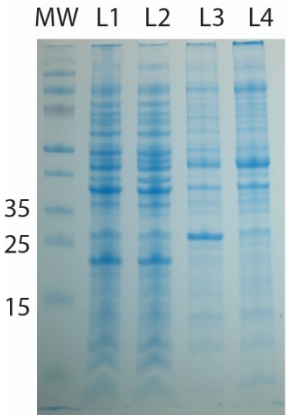


his-ROC 1323-1529:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

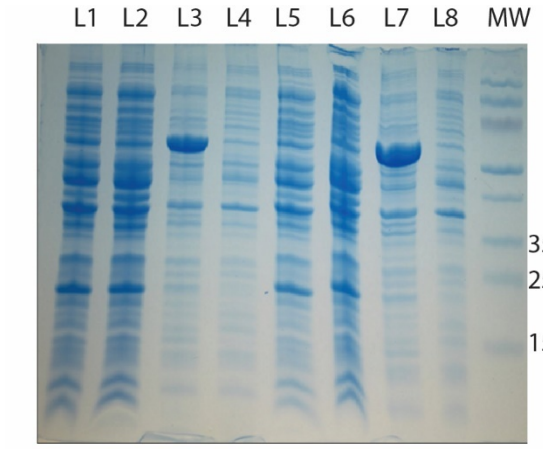
his-ROC 1320-1526:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced



his-ROC 1315-1529:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

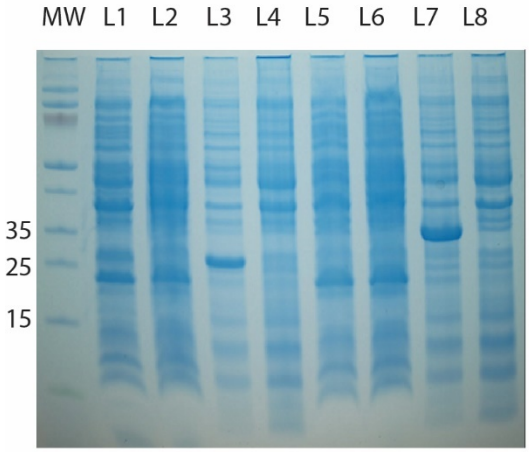


his-ROC 1329-1850:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

his-ROC 1329-1845:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

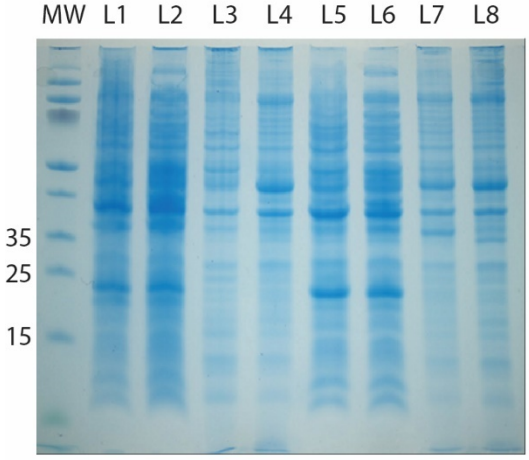


his-ROC 1329-1540:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

his-ROC 1329-1587:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

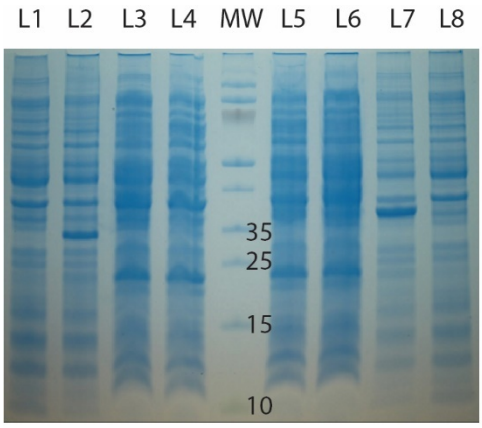


his-ROC 1276-1520:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

his-ROC 1276-1540:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

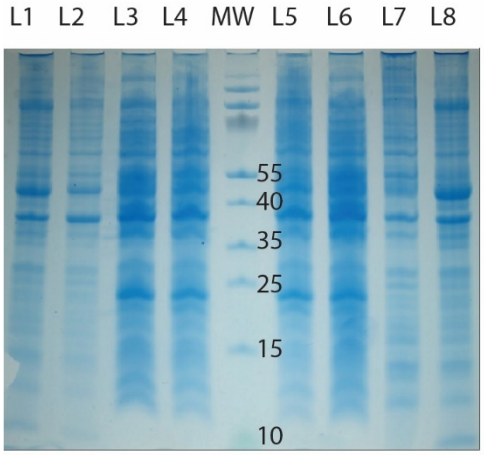


his-ROC 1301-1540:

L1: insoluble fraction, uninduced
L2: insoluble fraction, induced
L3: soluble fraction, uninduced
L4: soluble fraction. induced

his-ROC 1301-1587:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

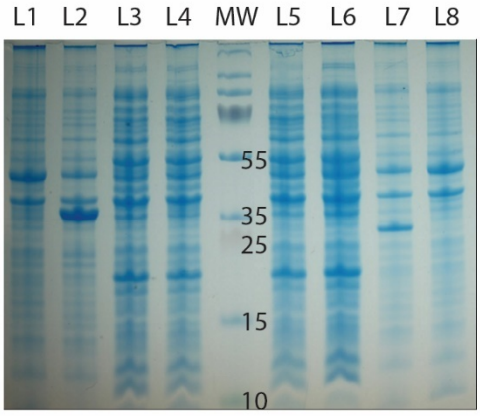


his-ROC 1276-1587:

L1: insoluble fraction, uninduced
L2: insoluble fraction, induced
L3: soluble fraction, uninduced
L4: soluble fraction. induced

his-ROC 1301-1520:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

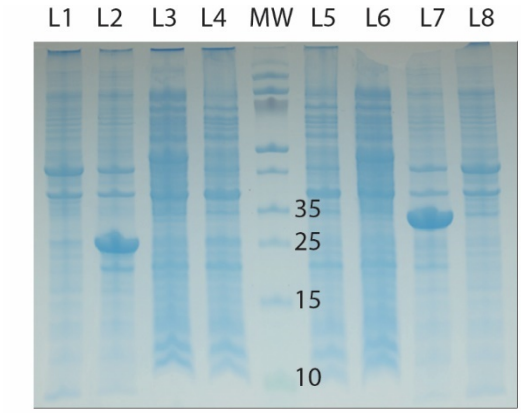


his-ROC 1301-1587:

L1: insoluble fraction, uninduced
L2: insoluble fraction, induced
L3: soluble fraction, uninduced
L4: soluble fraction. induced

his-ROC 1301-1540:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

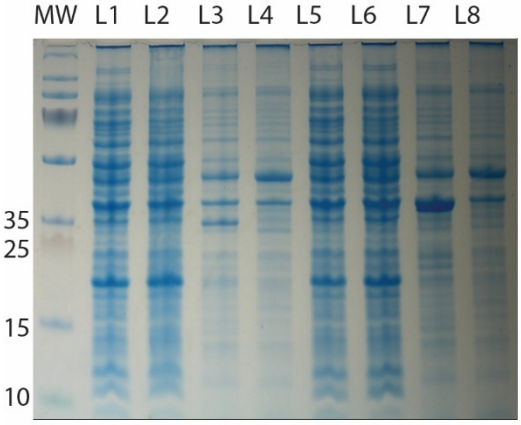


his-ROC 1320-1540:

L1: insoluble fraction, uninduced
L2: insoluble fraction, induced
L3: soluble fraction, uninduced
L4: soluble fraction. induced

his-ROC 1320-1587:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

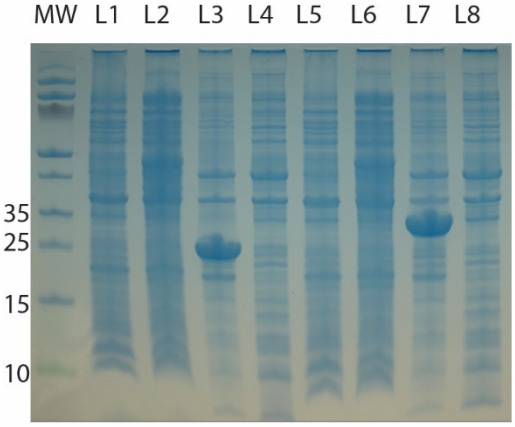


his-ROC 1276-1540:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

his-ROC 1276-1587:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

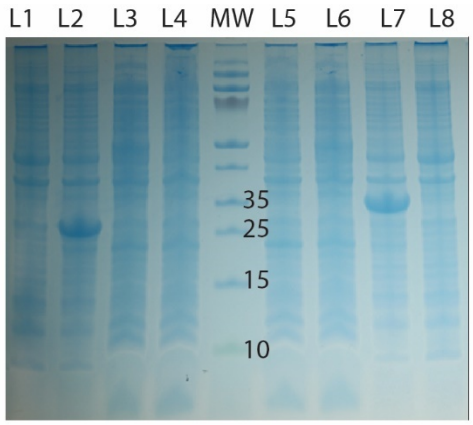


his-ROC 1315-1540:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

his-ROC 1315-1587:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

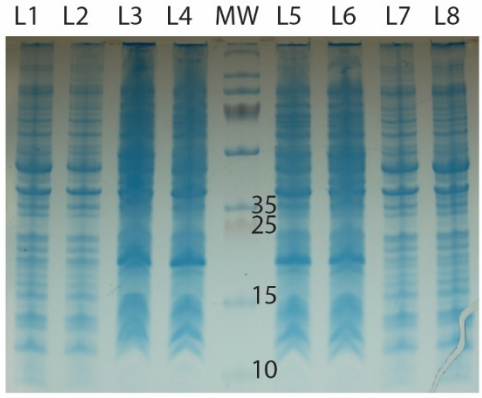


his-ROC 1323-1540:

L1: insoluble fraction, uninduced
L2: insoluble fraction, induced
L3: soluble fraction, uninduced
L4: soluble fraction. induced

his-ROC 1323-1587:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

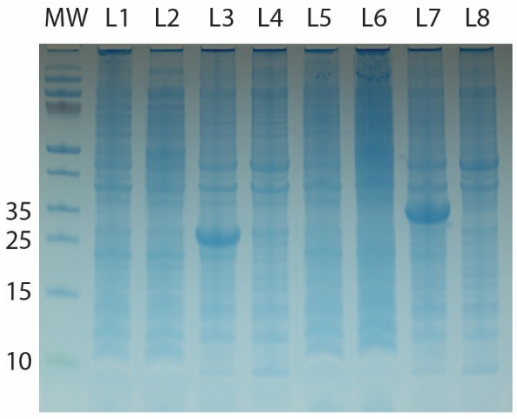


his-ROC 1261-1540:

L1: insoluble fraction, uninduced
L2: insoluble fraction, induced
L3: soluble fraction, uninduced
L4: soluble fraction. induced

his-ROC 1261-1580:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

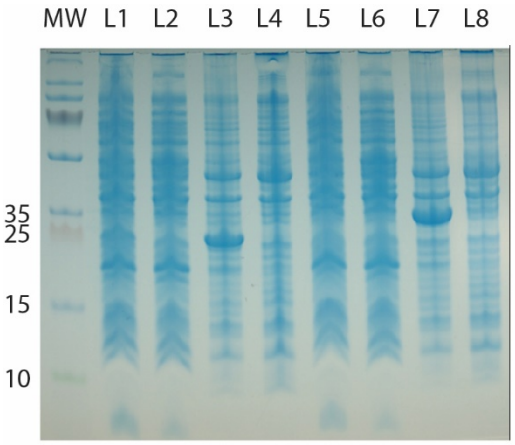


his-ROC 1326-1540:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

his-ROC 1326-1587:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

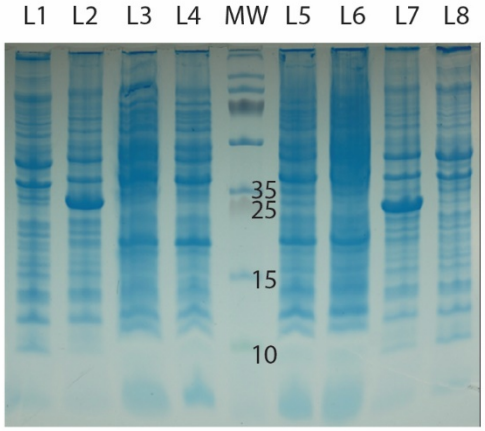


his-ROC 1312-1540:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

his-ROC 1312-1580:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

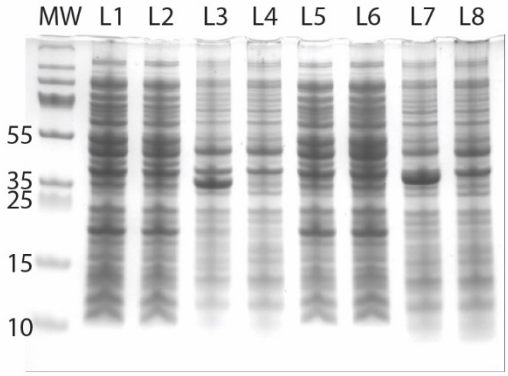


his-ROC 1320-1580:

L1: insoluble fraction, uninduced
L2: insoluble fraction, induced
L3: soluble fraction, uninduced
L4: soluble fraction. induced

his-ROC 1329-1580:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

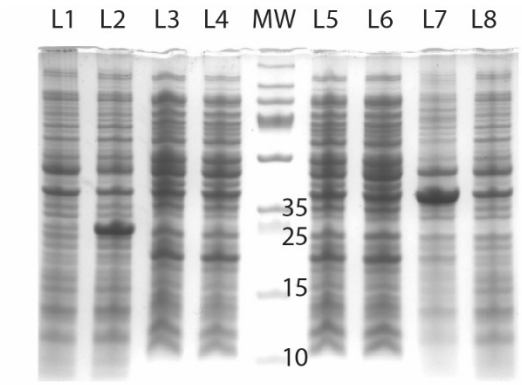


his-ROC 1256-1540:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

his-ROC 1256-1561:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

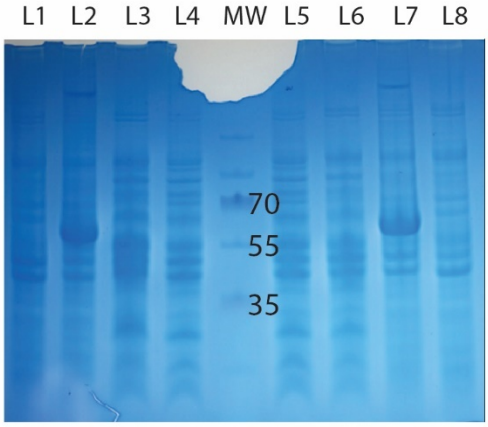


his-ROC 1329-1561:

L1: insoluble fraction, uninduced
L2: insoluble fraction, induced
L3: soluble fraction, uninduced
L4: soluble fraction. induced

his-ROC 1329-1643:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

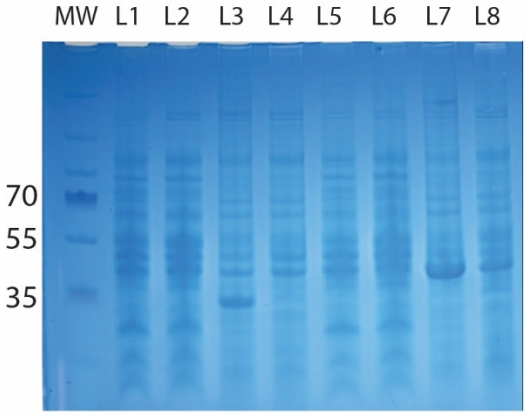


his-ROC 1329-1832:

L1: insoluble fraction, uninduced
L2: insoluble fraction, induced
L3: soluble fraction, uninduced
L4: soluble fraction. induced

his-ROC 1320-1832:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

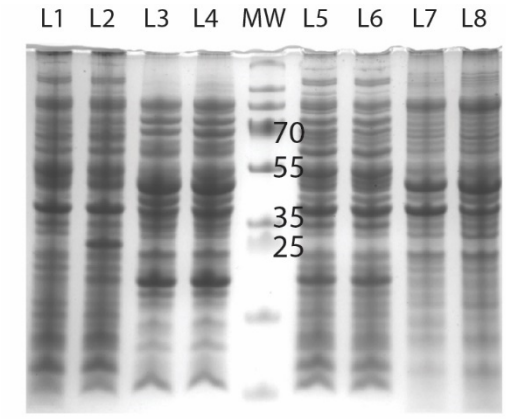


his-ROC 1320-1561:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

his-ROC 1320-1643:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

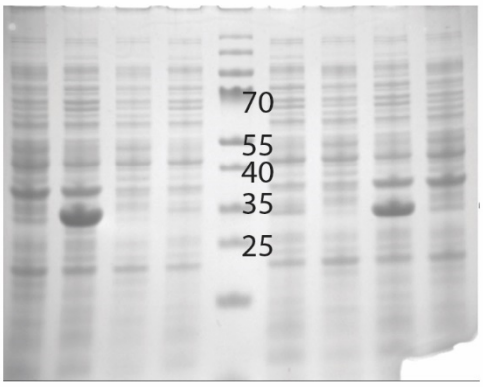


his-ROC 1307-1540:

L1: insoluble fraction, uninduced
L2: insoluble fraction, induced
L3: soluble fraction, uninduced
L4: soluble fraction. induced

his-ROC 1256-1580:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

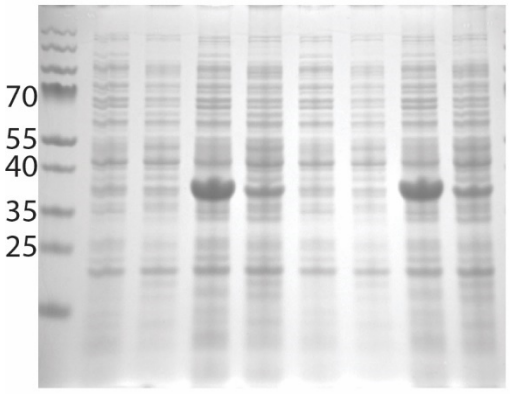


his-ROC 1329-1599:

L1: insoluble fraction, uninduced
L2: insoluble fraction, induced
L3: soluble fraction, uninduced
L4: soluble fraction. induced

his-ROC 1320-1580:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

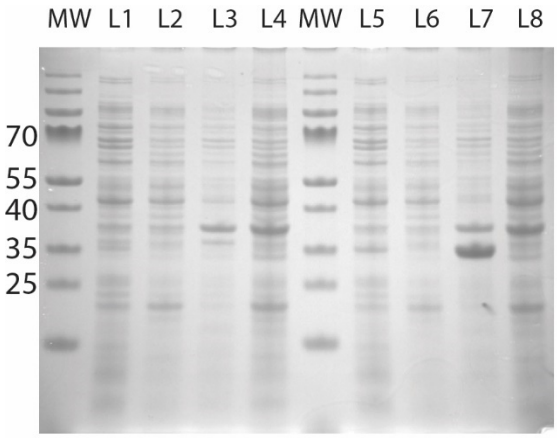


his-ROC 1329-1655:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

his-ROC 1329-1657:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

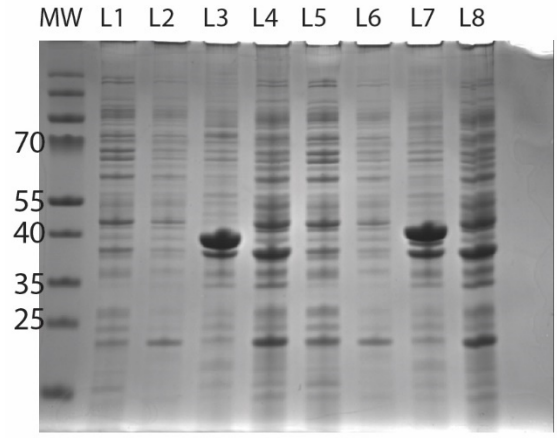


his-ROC 1301-1580:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

his-ROC 1312-1580:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

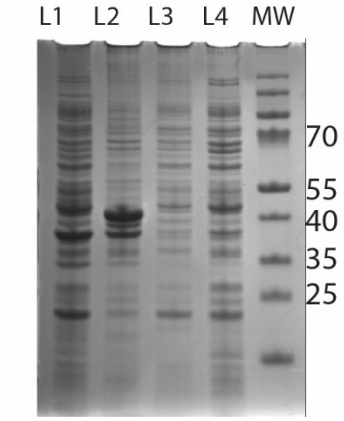


his-ROC 1312-1657:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

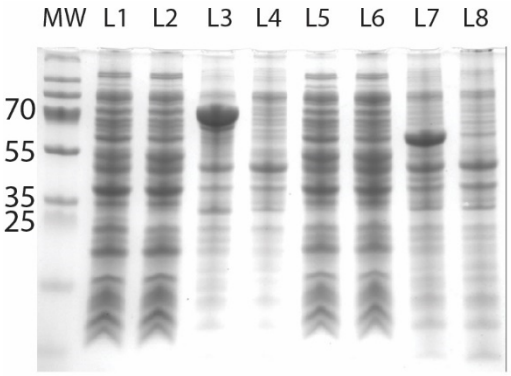
his-ROC 1312-1675:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced



his-ROC 1312-1672:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

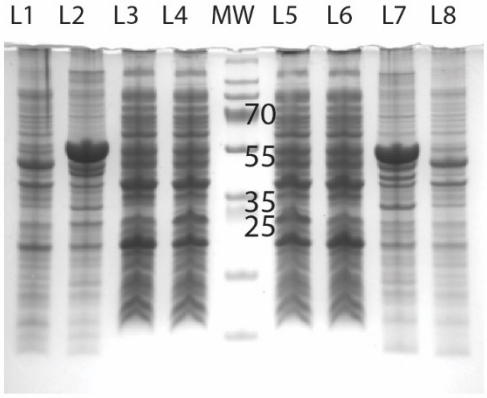


GST-ROC 1329-1832:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

his-ROC 1312-1832:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

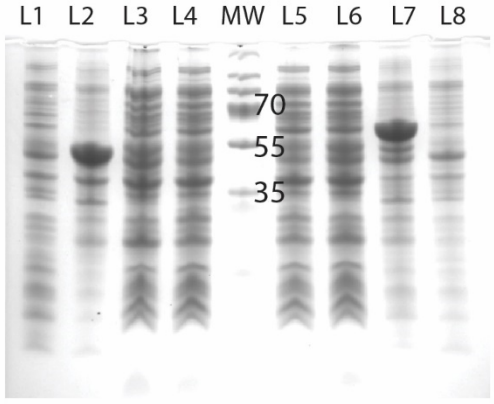


GST-ROC 1329-1580:

L1: insoluble fraction, uninduced
L2: insoluble fraction, induced
L3: soluble fraction, uninduced
L4: soluble fraction. induced

GST-ROC 1329-1561:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

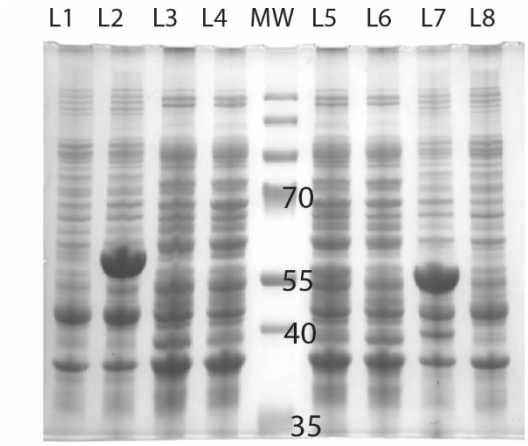


GST-ROC 1329-1540:

L1: insoluble fraction, uninduced
L2: insoluble fraction, induced
L3: soluble fraction, uninduced
L4: soluble fraction. induced

GST-ROC 1329-1643:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

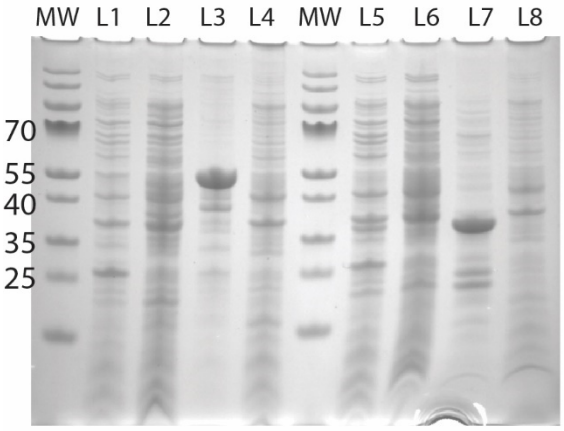


GST-ROC 1256-1540:

L1: insoluble fraction, uninduced
L2: insoluble fraction, induced
L3: soluble fraction, uninduced
L4: soluble fraction. induced

GST-ROC 1301-1540:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

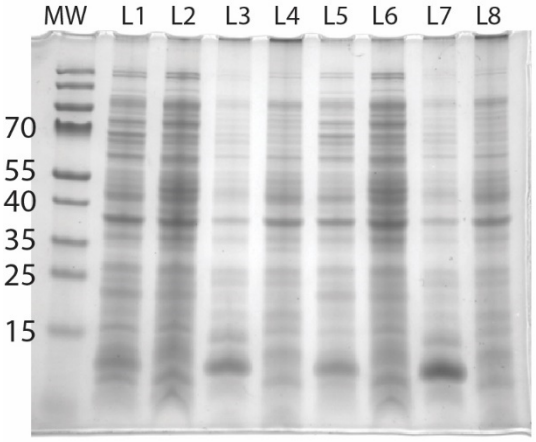


GST-ROC 1329-1580:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

GST-COR 1516-1618 (murine):

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

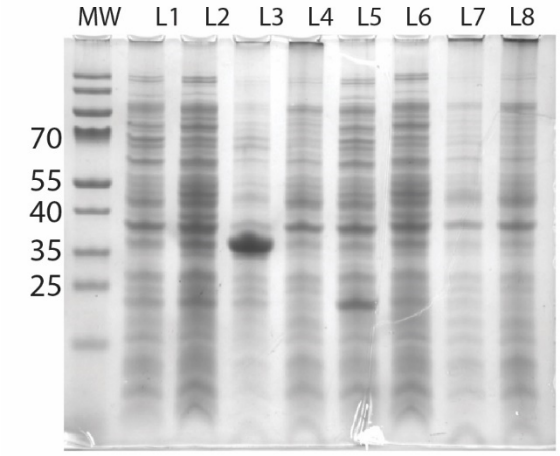


his-COR 1618-1832:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction. uninduced

GST-COR 1618-1832:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction. uninduced

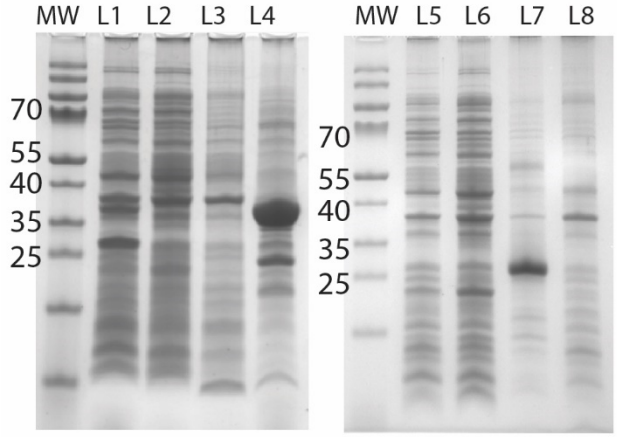


his-ROC 1329-1618:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction, uninduced

GST-ROC 1329-1618:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction, uninduced

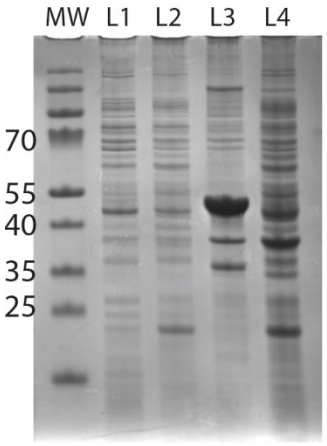


GST-ROC 1516-1618:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction, uninduced

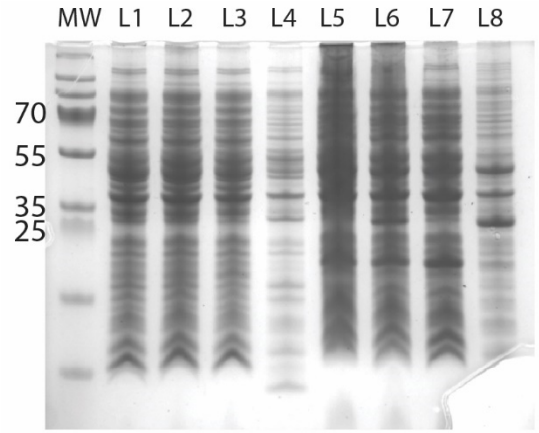
his-COR 1599-1832:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction, uninduced



his-sumo-ROC 1312-1580:

L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction, uninduced

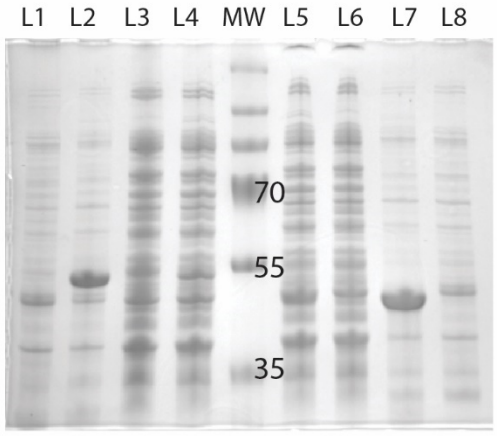


his-sumo-ROC 1329-1540:

L1: total lysate, uninduced
L2: total lysate, induced
L3: soluble fraction, induced
L4: insoluble fraction, induced

his-ROC 1312-1561:

L5: total lysate, uninduced
L6: total lysate, induced
L7: soluble fraction, induced
L8: insoluble fraction, induced

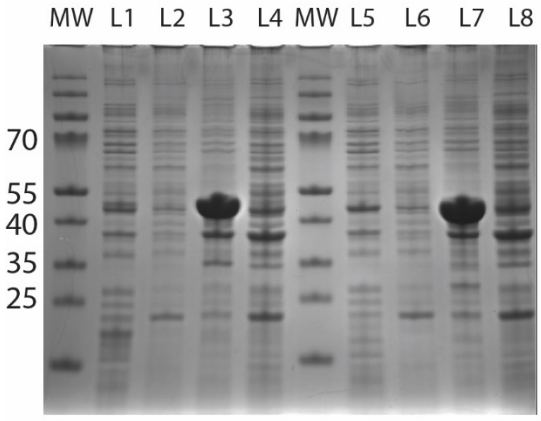


his-sumo-ROC 1256-1540:

L1: insoluble fraction, uninduced
L2: insoluble fraction, induced
L3: soluble fraction, uninduced
L4: soluble fraction, induced

his-sumo-ROC 1301-1540:

L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction, uninduced



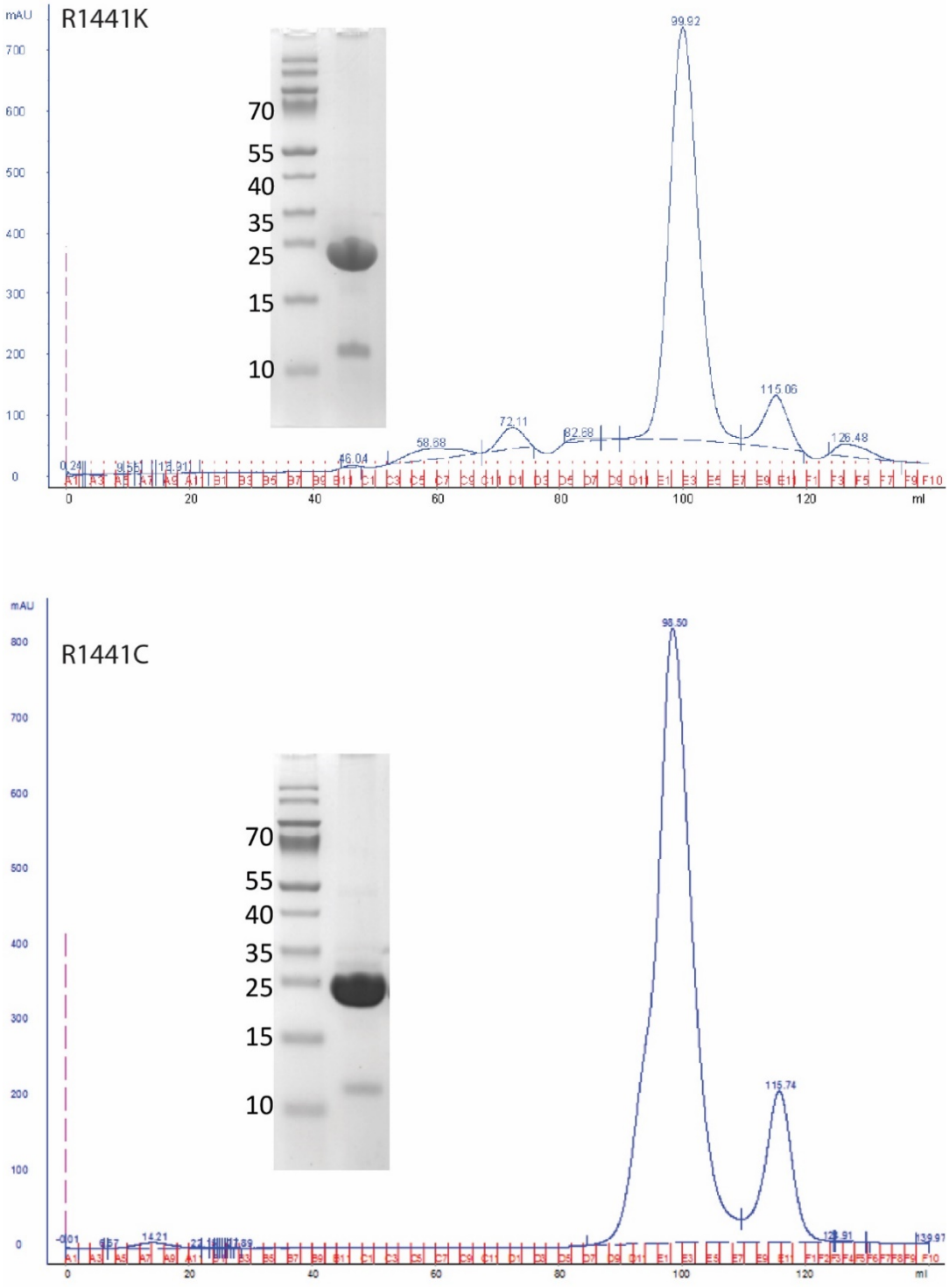
his-sumo-ROC 1320-1580:

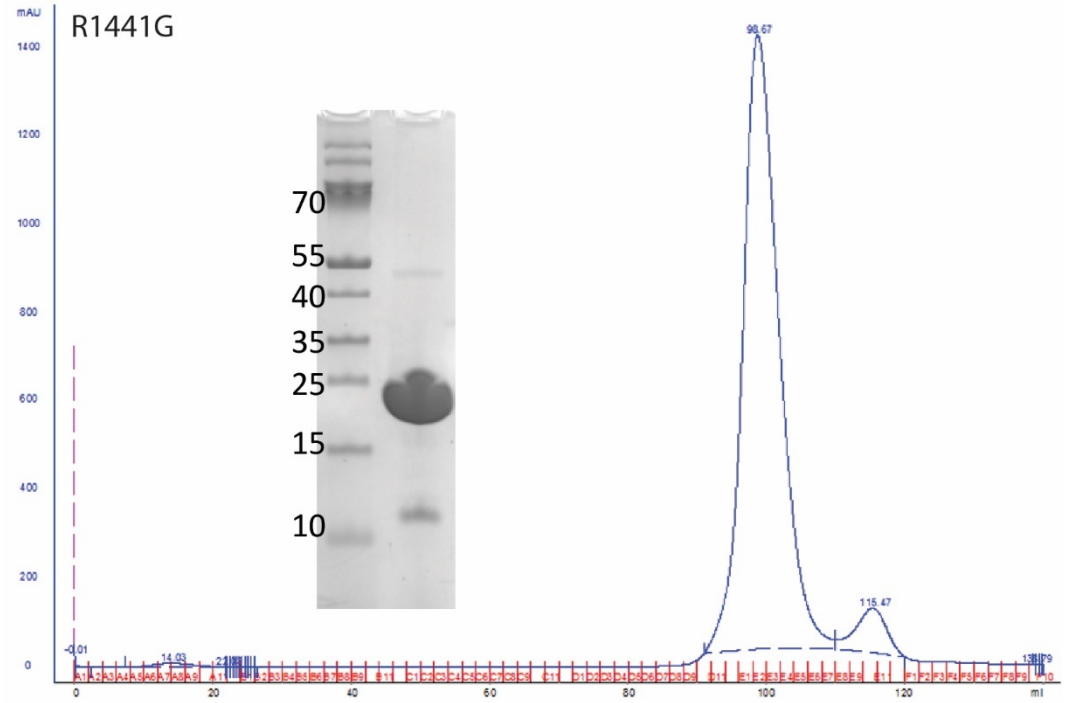
L1: soluble fraction, induced
L2: soluble fraction, uninduced
L3: insoluble fraction, induced
L4: insoluble fraction, uninduced

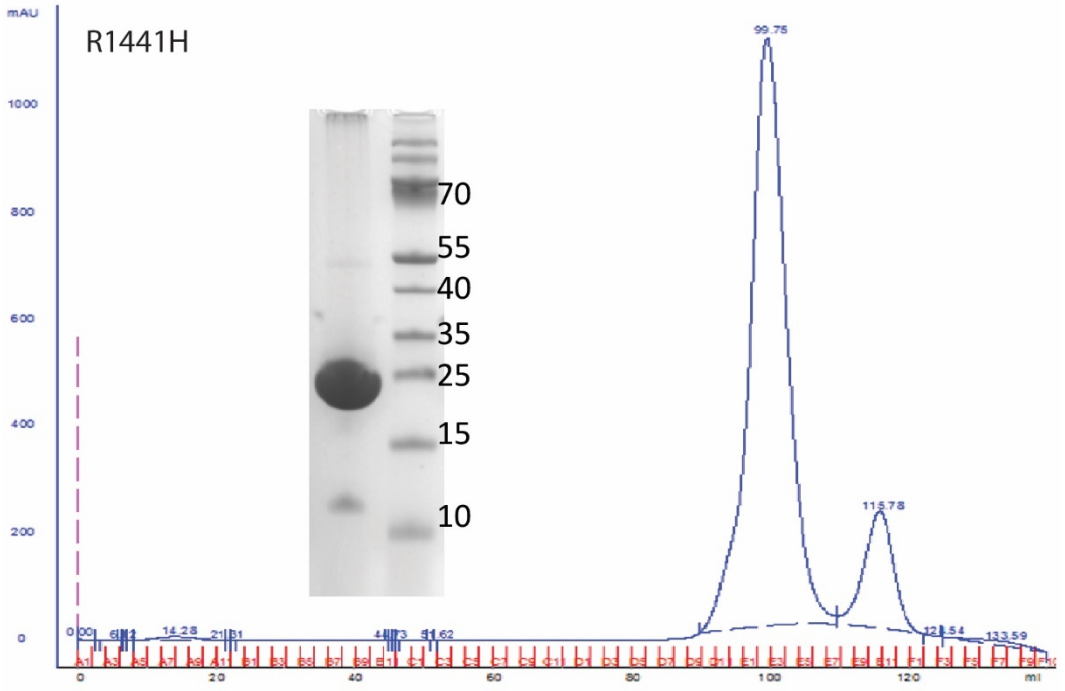
his-sumo-ROC 1329-1580:

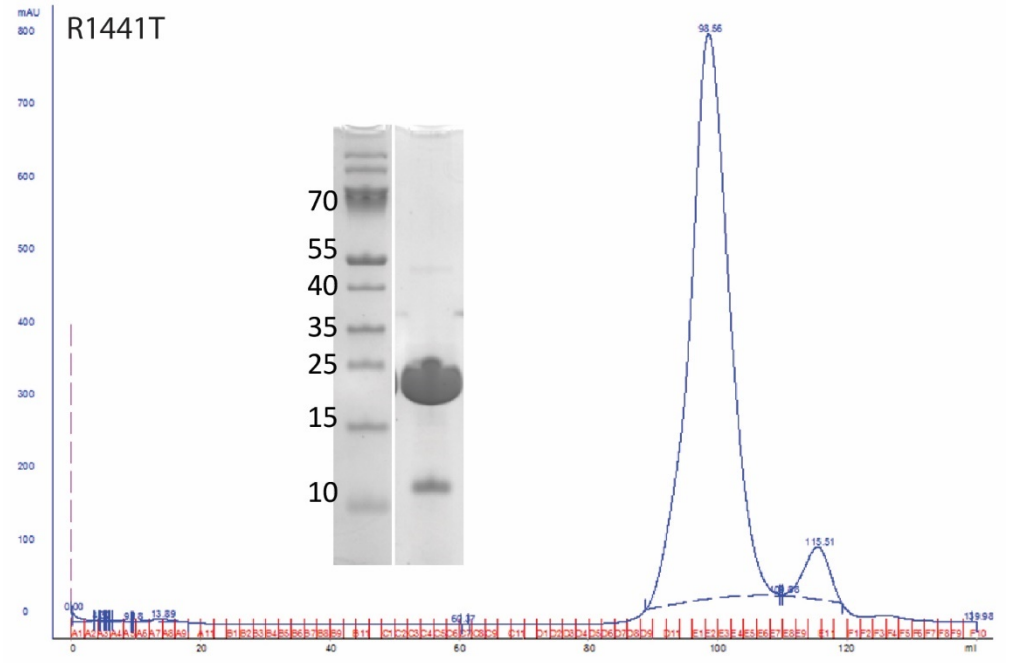
L5: soluble fraction, induced
L6: soluble fraction, uninduced
L7: insoluble fraction, induced
L8: insoluble fraction, uninduced

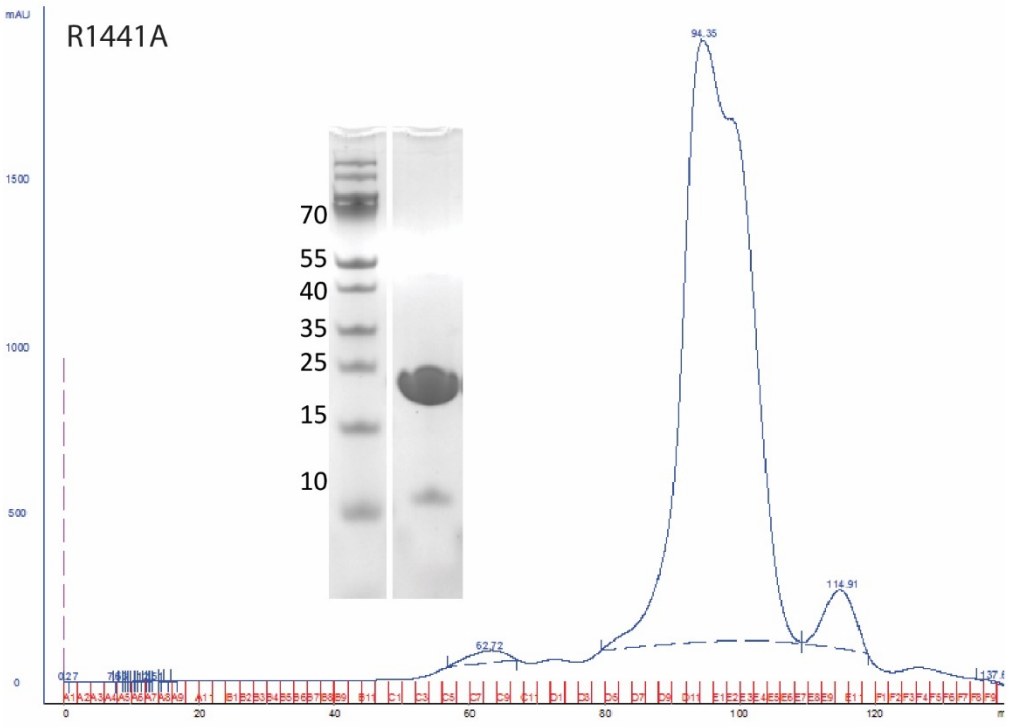
Figure A.29. SDS PAGE and size exclusion profiles of the purification of ROC with R1441 mutations

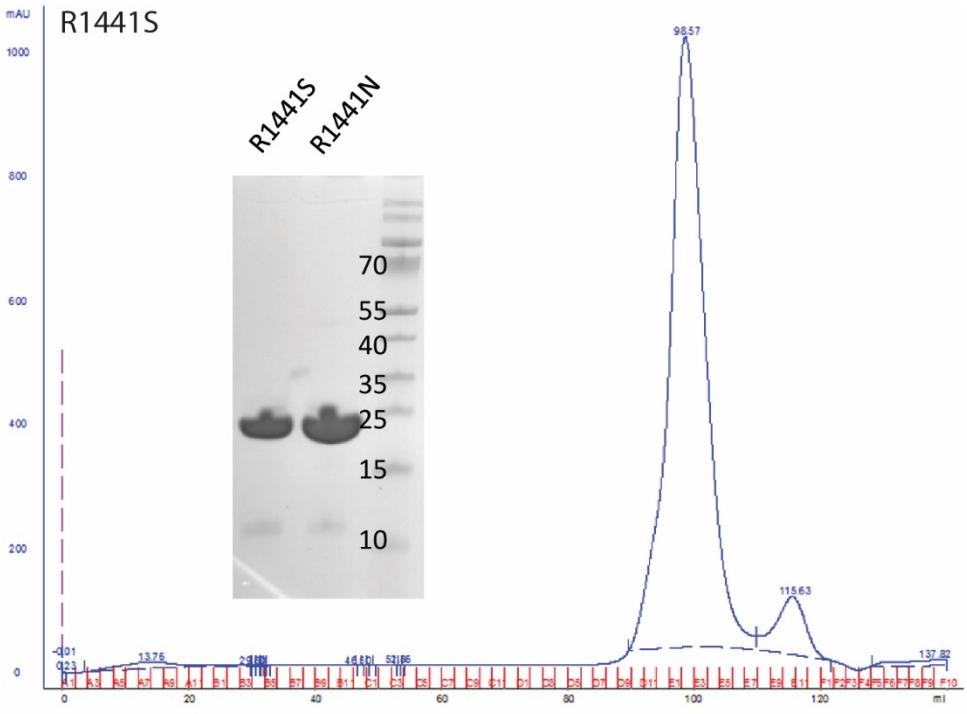


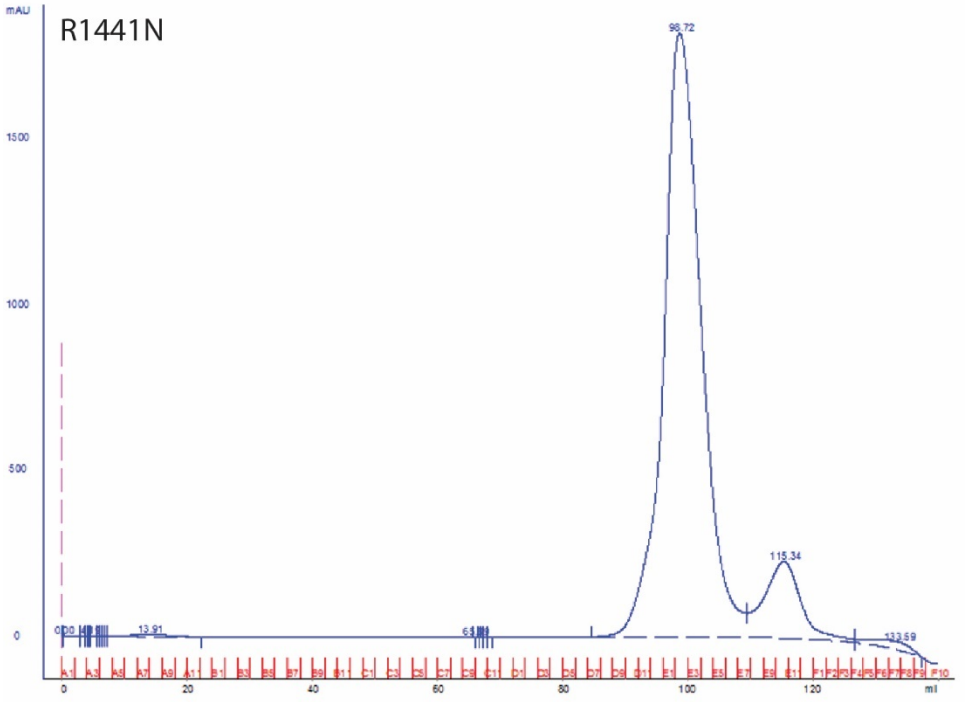

R1441G
mAU
98.67
115.47
70
55
40
35
25
15
10
ml

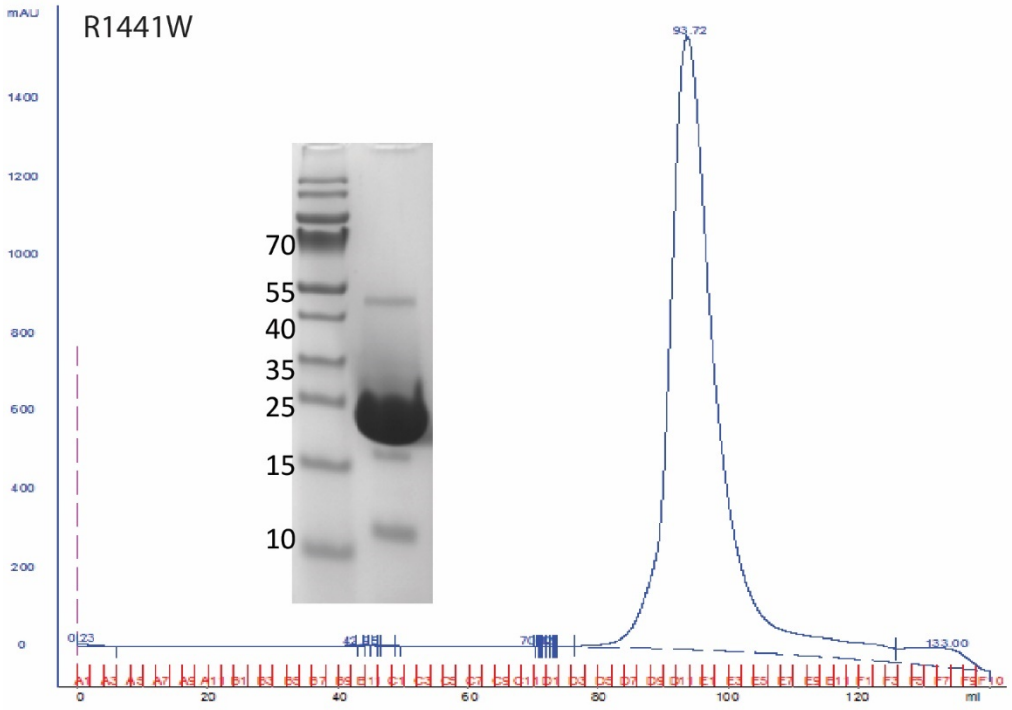

R1441H
mAU
99.75
115.78
70
55
40
35
25
15
10
ml


R1441T
mAU
800
700
600
500
400
300
200
100
0
70
55
40
35
25
15
10
98.56
115.51
0.00
13.89
59.17
139.98
0
20
40
60
80
100
120
ml


R1441A
mAU
1500
1000
500
0
70
55
40
35
25
15
10
94.35
114.91
62.72
0.27
0
20
40
60
80
100
120


R1441S
R1441S
R1441N
70
55
40
35
25
15
10
98.57
115.63
137.82
mAU
ml


R1441N
98.72
115.34
133.59
mAU
ml


R1441W
mAU
1400
1200
1000
800
600
400
200
0
93.72
70
55
40
35
25
15
10
133.00
0
20
40
60
80
100
120
ml

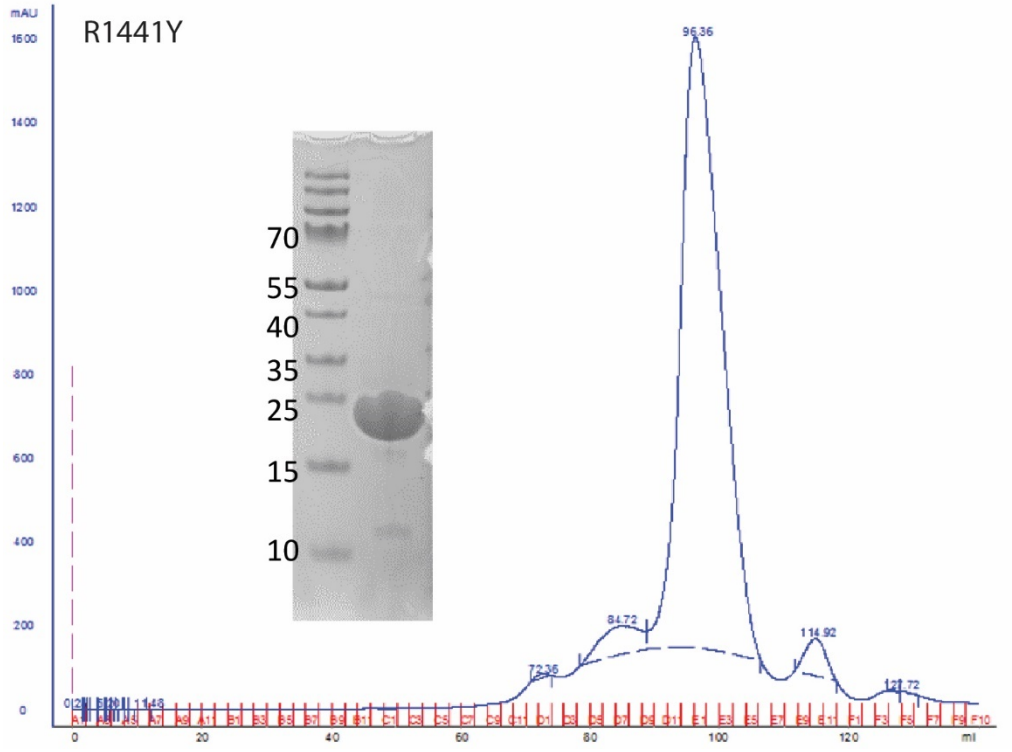
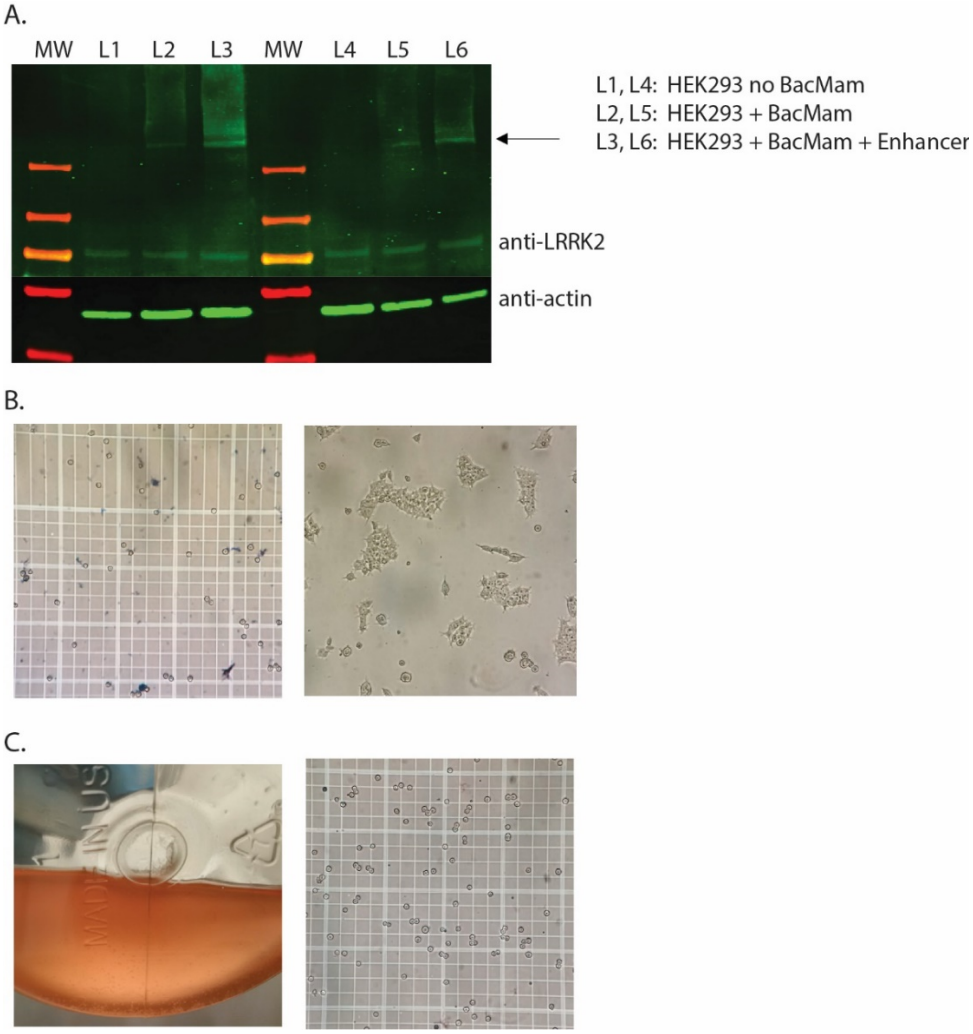

R1441Y
mAU
1600
1400
1200
1000
800
600
400
200
0
96.36
84.72
114.92
72.36
127.72
70
55
40
35
25
15
10
0
20
40
60
80
100
120
ml

Figure A.30. Effect of BacMam enhancer on LRRK2 expression in HEK293T cells and the reviving culture of Expi293 cells



A). Comparison of LRRK2 expression with or without BacMam Enhancer Solution (Thermo Fisher). **B).** Expi293 cells are revived in adhesion culture for 2 passages before adaption into suspension culture (right), because cell viability is bad when directly revived into suspension culture (left). **C).** Adaptation of Expi293 cells from adhesion culture (containing FBS) to serum free suspension culture with good viability (right). Cells formed some observable clusters in the first two passages of suspension culture (left), which is normal. Cells would become fully suspended after three passages.

References

Adams, P. D., P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart (2010). "PHENIX: a comprehensive Python-based system for macromolecular structure solution." Acta Crystallogr D Biol Crystallogr **66**(Pt 2): 213-221.

Alegre-Abarrategui, J., H. Christian, M. M. Lufino, R. Mutihac, L. L. Venda, O. Ansorge and R. Wade-Martins (2009). "LRRK2 regulates autophagic activity and localizes to specific membrane microdomains in a novel human genomic reporter cellular model." Hum Mol Genet **18**(21): 4022-4034.

Anand, V. S., L. J. Reichling, K. Lipinski, W. Stochaj, W. Duan, K. Kelleher, P. Pungaliya, E. L. Brown, P. H. Reinhart, R. Somberg, W. D. Hirst, S. M. Riddle and S. P. Braithwaite (2009). "Investigation of leucine-rich repeat kinase 2 : enzymological properties and novel assays." FEBS J **276**(2): 466-478.

Andrade, M. A. and P. Bork (1995). "HEAT repeats in the Huntington's disease protein." Nat Genet **11**(2): 115-116.

Andrade, M. A., C. Petosa, S. I. O'Donoghue, C. W. Muller and P. Bork (2001). "Comparison of ARM and HEAT protein repeats." J Mol Biol **309**(1): 1-18.

Baptista, M. A., K. D. Dave, M. A. Frasier, T. B. Sherer, M. Greeley, M. J. Beck, J. S. Varsho, G. A. Parker, C. Moore, M. J. Churchill, C. K. Meshul and B. K. Fiske (2013). "Loss of leucine-rich repeat kinase 2 (LRRK2) in rats leads to progressive abnormal phenotypes in peripheral organs." PLoS One **8**(11): e80705.

Beilina, A., I. N. Rudenko, A. Kaganovich, L. Civiero, H. Chau, S. K. Kalia, L. V. Kalia, E. Lobbestael, R. Chia, K. Ndukwe, J. Ding, M. A. Nalls, C. International Parkinson's Disease Genomics, C. North American Brain Expression, M. Olszewski, D. N. Hauser, R. Kumaran, A. M. Lozano, V. Baekelandt, L. E. Greene, J. M. Taymans, E. Greggio and M. R. Cookson (2014). "Unbiased screen for interactors of leucine-rich repeat kinase 2 supports a common pathway for sporadic and familial Parkinson disease." Proc Natl Acad Sci U S A **111**(7): 2626-2631.

Bella, J., K. L. Hindle, P. A. McEwan and S. C. Lovell (2008). "The leucine-rich repeat structure." Cell Mol Life Sci **65**(15): 2307-2333.

Berger, Z., K. A. Smith and M. J. Lavoie (2010). "Membrane localization of LRRK2 is associated with increased formation of the highly active LRRK2 dimer and changes in its phosphorylation." Biochemistry **49**(26): 5511-5523.

Biosa, A., A. Trancikova, L. Civiero, L. Glauser, L. Bubacco, E. Greggio and D. J. Moore (2013). "GTPase activity regulates kinase activity and cellular phenotypes of Parkinson's disease-associated LRRK2." Hum Mol Genet **22**(6): 1140-1156.

Biskup, S., D. J. Moore, F. Celsi, S. Higashi, A. B. West, S. A. Andrabi, K. Kurkinen, S. W. Yu, J. M. Savitt, H. J. Waldvogel, R. L. Faull, P. C. Emson, R. Torp, O. P. Ottersen, T. M. Dawson and V. L. Dawson (2006). "Localization of LRRK2 to membranous and vesicular structures in mammalian brain." Ann Neurol **60**(5): 557-569.

Bosgraaf, L. and P. J. Van Haastert (2003). "Roc, a Ras/GTPase domain in complex proteins." Biochim Biophys Acta **1643**(1-3): 5-10.

Brooks, B. R., C. L. Brooks, 3rd, A. D. Mackerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R.

Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York and M. Karplus (2009). "CHARMM: the biomolecular simulation program." J Comput Chem **30**(10): 1545-1614.

Chakrabarty, B. and N. Parekh (2014). "Identifying tandem Ankyrin repeats in protein structures." BMC Bioinformatics **15**: 6599.

Chavas, L. M. G., S. Torii, H. Kamikubo, M. Kawasaki, K. Ihara, R. Kato, M. Kataoka, T. Izumi and S. Wakatsuki (2007). "Structure of the small GTPase Rab27b shows an unexpected swapped dimer." Acta Crystallographica Section D **63**(7): 769-779.

Chinta, S. J. and J. K. Andersen (2005). "Dopaminergic neurons." Int J Biochem Cell Biol **37**(5): 942-946.

Chook, Y. M. and G. Blobel (1999). "Structure of the nuclear transport complex karyopherin-beta2-Ran x GppNHp." Nature **399**(6733): 230-237.

Civiero, L., S. Dihanich, P. A. Lewis and E. Greggio (2014). "Genetic, structural, and molecular insights into the function of ras of complex proteins domains." Chem Biol **21**(7): 809-818.

Cookson, M. R. (2010). "The role of leucine-rich repeat kinase 2 (LRRK2) in Parkinson's disease." Nat Rev Neurosci **11**(12): 791-797.

Cremades, N., S. I. Cohen, E. Deas, A. Y. Abramov, A. Y. Chen, A. Orte, M. Sandal, R. W. Clarke, P. Dunne, F. A. Aprile, C. W. Bertoncini, N. W. Wood, T. P. Knowles, C. M. Dobson and D. Klenerman (2012). "Direct Observation of the Interconversion of Normal and Toxic Forms of alpha-Synuclein." Cell **149**(5): 1048-1059.

Cuervo, A. M. and E. Wong (2014). "Chaperone-mediated autophagy: roles in disease and aging." Cell Res **24**(1): 92-104.

Dachsel, J. C., J. P. Taylor, S. S. Mok, O. A. Ross, K. M. Hinkle, R. M. Bailey, J. H. Hines, J. Szutu, B. Madden, L. Petrucelli and M. J. Farrer (2007). "Identification of potential protein interactors of Lrrk2." Parkinsonism Relat Disord **13**(7): 382-385.

Daniels, V., R. Vancraenenbroeck, B. M. Law, E. Greggio, E. Lobbestael, F. Gao, M. De Maeyer, M. R. Cookson, K. Harvey, V. Baekelandt and J. M. Taymans (2011). "Insight into the mode of action of the LRRK2 Y1699C pathogenic mutant." J Neurochem **116**(2): 304-315.

Deng, J., P. A. Lewis, E. Greggio, E. Sluch, A. Beilina and M. R. Cookson (2008). "Structure of the ROC domain from the Parkinson's disease-associated leucine-rich repeat kinase 2 reveals a dimeric GTPase." Proc Natl Acad Sci U S A **105**(5): 1499-1504.

Deyaert, E., L. Wauters, G. Guaitoli, A. Konijnenberg, M. Leemans, S. Terheyden, A. Petrovic, R. Gallardo, L. M. Nederveen-Schippers, P. S. Athanasopoulos, H. Pots, P. J. M. Van Haastert, F. Sobott, C. J. Gloeckner, R. Efremov, A. Kortholt and W. Versees (2017). "A homologue of the Parkinson's disease-associated protein LRRK2 undergoes a monomer-dimer transition during GTP turnover." Nat Commun **8**(1): 1008.

Eisenberg, E. and L. E. Greene (2007). "Multiple roles of auxilin and hsc70 in clathrin-mediated endocytosis." Traffic **8**(6): 640-646.

Elbaz, A. and C. Tranchant (2007). "Epidemiologic studies of environmental exposures in Parkinson's disease." J Neurol Sci **262**(1-2): 37-44.

Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallogr D Biol Crystallogr **60**(Pt 12 Pt 1): 2126-2132.

Emsley, P., B. Lohkamp, W. G. Scott and K. Cowtan (2010). "Features and development of Coot." Acta Crystallogr D Biol Crystallogr **66**(Pt 4): 486-501.

Fiser, A. and A. Sali (2003). "Modeller: generation and refinement of homology-based protein structure models." Methods Enzymol **374**: 461-491.

Fleming, L., J. B. Mann, J. Bean, T. Briggle and J. R. Sanchez-Ramos (1994). "Parkinson's disease and brain levels of organochlorine pesticides." Ann Neurol **36**(1): 100-103.

Funayama, M., K. Hasegawa, H. Kowa, M. Saito, S. Tsuji and F. Obata (2002). "A new locus for Parkinson's disease (PARK8) maps to chromosome 12p11.2-q13.1." Ann Neurol **51**(3): 296-301.

Gasper, R., S. Meyer, K. Gotthardt, M. Sirajuddin and A. Wittinghofer (2009). "It takes two to tango: regulation of G proteins by dimerization." Nat Rev Mol Cell Biol **10**(6): 423-429.

Gilsbach, B. K., F. Y. Ho, I. R. Vetter, P. J. van Haastert, A. Wittinghofer and A. Kortholt (2012). "Roco kinase structures give insights into the mechanism of Parkinson disease-related leucine-rich-repeat kinase 2 mutations." Proc Natl Acad Sci U S A **109**(26): 10322-10327.

Glick, D., S. Barth and K. F. Macleod (2010). "Autophagy: cellular and molecular mechanisms." J Pathol **221**(1): 3-12.

Goedert, M. (2001). "Alpha-synuclein and neurodegenerative diseases." Nat Rev Neurosci **2**(7): 492-501.

Goedert, M., M. G. Spillantini, K. Del Tredici and H. Braak (2013). "100 years of Lewy pathology." Nat Rev Neurol **9**(1): 13-24.

Goldfarb, D. S., A. H. Corbett, D. A. Mason, M. T. Harreman and S. A. Adam (2004). "Importin alpha: a multipurpose nuclear-transport receptor." Trends Cell Biol **14**(9): 505-514.

Gorostidi, A., J. Ruiz-Martinez, A. Lopez de Munain, A. Alzualde and J. F. Marti Masso (2009). "LRRK2 G2019S and R1441G mutations associated with Parkinson's disease are common in the Basque Country, but relative prevalence is determined by ethnicity." Neurogenetics **10**(2): 157-159.

Gotthardt, K., M. Weyand, A. Kortholt, P. J. Van Haastert and A. Wittinghofer (2008). "Structure of the Roc-COR domain tandem of C. tepidum, a prokaryotic homologue of the human LRRK2 Parkinson kinase." EMBO J **27**(16): 2239-2249.

Greggio, E., S. Jain, A. Kingsbury, R. Bandopadhyay, P. Lewis, A. Kaganovich, M. P. van der Brug, A. Beilina, J. Blackinton, K. J. Thomas, R. Ahmad, D. W. Miller, S. Kesavapany, A. Singleton, A. Lees, R. J. Harvey, K. Harvey and M. R. Cookson (2006). "Kinase activity is required for the toxic effects of mutant LRRK2/dardarin." Neurobiol Dis **23**(2): 329-341.

Greggio, E., I. Zambrano, A. Kaganovich, A. Beilina, J. M. Taymans, V. Daniels, P. Lewis, S. Jain, J. Ding, A. Syed, K. J. Thomas, V. Baekelandt and M. R. Cookson (2008). "The Parkinson disease-associated leucine-rich repeat kinase 2 (LRRK2) is a dimer that undergoes intramolecular autophosphorylation." J Biol Chem **283**(24): 16906-16914.

Guo, L., P. N. Gandhi, W. Wang, R. B. Petersen, A. L. Wilson-Delfosse and S. G. Chen (2007). "The Parkinson's disease-associated protein, leucine-rich repeat kinase 2 (LRRK2), is an authentic GTPase that stimulates kinase activity." Exp Cell Res **313**(16): 3658-3670.

Helip-Wooley, A. and J. G. Thoene (2004). "Sucrose-induced vacuolation results in increased expression of cholesterol biosynthesis and lysosomal genes." Exp Cell Res **292**(1): 89-100.

Hernan, M. A., B. Takkouche, F. Caamano-Isorna and J. J. Gestal-Otero (2002). "A meta-analysis of coffee drinking, cigarette smoking, and the risk of Parkinson's disease." Ann Neurol **52**(3): 276-284.

Ito, G., T. Okai, G. Fujino, K. Takeda, H. Ichijo, T. Katada and T. Iwatsubo (2007). "GTP binding is essential to the protein kinase activity of LRRK2, a causative gene product for familial Parkinson's disease." Biochemistry **46**(5): 1380-1388.

James, N. G., M. A. Digman, E. Gratton, B. Barylko, X. Ding, J. P. Albanesi, M. S. Goldberg and D. M. Jameson (2012). "Number and brightness analysis of LRRK2 oligomerization in live cells." Biophys J **102**(11): L41-43.

Jean-Marc Taymans, Mark R. Cookson (2011). "LRRK2 Kinase Activity is Dependent on LRRK2 GTP Binding Capacity but Independent of LRRK2 GTP Binding." PLoS One **6**(8).

Jovic, M., M. Sharma, J. Rahajeng and S. Caplan (2010). "The early endosome: a busy sorting station for proteins at the crossroads." Histol Histopathol **25**(1): 99-112.

Kasper, D. L. (2015). Harrison's principles of internal medicine. New York, McGraw Hill Education.

Kempster, P. A., B. Hurwitz and A. J. Lees (2007). "A new look at James Parkinson's Essay on the Shaking Palsy." Neurology **69**(5): 482-485.

Klein, C. and A. Westenberger (2012). "Genetics of Parkinson's disease." Cold Spring Harb Perspect Med **2**(1): a008888.

Klein, C. L., G. Rovelli, W. Springer, C. Schall, T. Gasser and P. J. Kahle (2009). "Homo- and heterodimerization of ROCO kinases: LRRK2 kinase inhibition by the LRRK2 ROCO fragment." J Neurochem **111**(3): 703-715.

Kukimoto-Niino, M., A. Sakamoto, E. Kanno, K. Hanawa-Suetsugu, T. Terada, M. Shirouzu, M. Fukuda and S. Yokoyama (2008). "Structural basis for the exclusive specificity of Slac2-a/melanophilin for the Rab27 GTPases." Structure **16**(10): 1478-1490.

Langston, J. W., I. Irwin, E. B. Langston and L. S. Forno (1984). "1-Methyl-4-phenylpyridinium ion (MPP+): identification of a metabolite of MPTP, a toxin selective to the substantia nigra." Neurosci Lett **48**(1): 87-92.

Lees, A. J., J. Hardy and T. Revesz (2009). "Parkinson's disease." Lancet **373**(9680): 2055-2066.

Lesage, S., A. Durr, M. Tazir, E. Lohmann, A. L. Leutenegger, S. Janin, P. Pollak, A. Brice and G. French Parkinson's Disease Genetics Study (2006). "LRRK2 G2019S as a cause of Parkinson's disease in North African Arabs." N Engl J Med **354**(4): 422-423.

Lewis, P. A. (2009). "The function of ROCO proteins in health and disease." Biol Cell **101**(3): 183-191.

Lewis, P. A., E. Greggio, A. Beilina, S. Jain, A. Baker and M. R. Cookson (2007). "The R1441C mutation of LRRK2 disrupts GTP hydrolysis." Biochem Biophys Res Commun **357**(3): 668-671.

Li, X., Y. C. Tan, S. Poulose, C. W. Olanow, X. Y. Huang and Z. Yue (2007). "Leucine-rich repeat kinase 2 (LRRK2)/PARK8 possesses GTPase activity that is altered in familial Parkinson's disease R1441C/G mutants." J Neurochem **103**(1): 238-247.

Liu, M., B. Dobson, M. A. Glicksman, Z. Yue and R. L. Stein (2008). "Kinetic mechanistic studies of wild-type leucine-rich repeat kinase 2: characterization of the kinase and GTPase activities." Biochemistry **49**(9): 2008-2017.

Luk, K. C. and V. M. Lee (2014). "Modeling Lewy pathology propagation in Parkinson's disease." Parkinsonism Relat Disord **20 Suppl 1**: S85-87.

Luzon-Toro, B., E. Rubio de la Torre, A. Delgado, J. Perez-Tur and S. Hilfiker (2007). "Mechanistic insight into the dominant mode of the Parkinson's disease-associated G2019S LRRK2 mutation." Hum Mol Genet **16**(17): 2031-2039.

MacLeod, D. A., H. Rhinn, T. Kuwahara, A. Zolin, G. Di Paolo, B. D. McCabe, K. S. Marder, L. S. Honig, L. N. Clark, S. A. Small and A. Abeliovich (2013). "RAB7L1 interacts with LRRK2 to modify intraneuronal protein sorting and Parkinson's disease risk." Neuron **77**(3): 425-439.

Maekawa, T., M. Kubo, I. Yokoyama, E. Ohta and F. Obata (2010). "Age-dependent and cell-population-restricted LRRK2 expression in normal mouse spleen." Biochem Biophys Res Commun **392**(3): 431-435.

Malumbres, M. and M. Barbacid (2003). "RAS oncogenes: the first 30 years." Nat Rev Cancer **3**(6): 459-465.

Marin, I., W. N. van Egmond and P. J. van Haastert (2008). "The Roco protein family: a functional perspective." FASEB J **22**(9): 3103-3110.

Mata, I. F., M. Y. Davis, A. N. Lopez, M. O. Dorschner, E. Martinez, D. Yearout, B. A. Cholerton, S. C. Hu, K. L. Edwards, T. D. Bird and C. P. Zabetian (2016). "The discovery of LRRK2 p.R1441S, a novel mutation for Parkinson's disease, adds to the complexity of a mutational hotspot." Am J Med Genet B Neuropsychiatr Genet **171**(7): 925-930.

Mata, I. F., M. Y. Davis, A. N. Lopez, M. O. Dorschner, E. Martinez, D. Yearout, B. A. Cholerton, S. C. Hu, K. L. Edwards, T. D. Bird and C. P. Zabetian (2017). "The discovery of LRRK2 p.R1441S, a novel mutation for Parkinson's disease, adds to the complexity of a mutational hotspot." Am J Med Genet B Neuropsychiatr Genet **174**(1): 113.

Murshudov, G. N., P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long and A. A. Vagin (2011). "REFMAC5 for the refinement of macromolecular crystal structures." Acta Crystallogr D Biol Crystallogr **67**(Pt 4): 355-367.

Nandipati, S. and I. Litvan (2016). "Environmental Exposures and Parkinson's Disease." Int J Environ Res Public Health **13**(9).

Ness, D., Z. Ren, S. Gardai, D. Sharpnack, V. J. Johnson, R. J. Brennan, E. F. Brigham and A. J. Olaharski (2013). "Leucine-rich repeat kinase 2 (LRRK2)-deficient rats exhibit renal tubule injury and perturbations in metabolic and immunological homeostasis." PLoS One **8**(6): e66164.

Otwinowski, Z. and W. Minor (1997). "Processing of X-ray diffraction data collected in oscillation mode." Methods Enzymol **276**: 307-326.

Ozelius, L. J., G. Senthil, R. Saunders-Pullman, E. Ohmann, A. Deligtisch, M. Tagliati, A. L. Hunt, C. Klein, B. Henick, S. M. Hailpern, R. B. Lipton, J. Soto-Valencia, N. Risch and S. B. Bressman (2006). "LRRK2 G2019S as a cause of Parkinson's disease in Ashkenazi Jews." N Engl J Med **354**(4): 424-425.

Paisan-Ruiz, C., S. Jain, E. W. Evans, W. P. Gilks, J. Simon, M. van der Brug, A. Lopez de Munain, S. Aparicio, A. M. Gil, N. Khan, J. Johnson, J. R. Martinez, D. Nicholl,

I. M. Carrera, A. S. Pena, R. de Silva, A. Lees, J. F. Marti-Masso, J. Perez-Tur, N. W. Wood and A. B. Singleton (2004). "Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease." Neuron **44**(4): 595-600.

Petsko, G. A. (2006). "The next epidemic." Genome Biol **7**(5): 108.

Ramonet, D., J. P. Daher, B. M. Lin, K. Stafa, J. Kim, R. Banerjee, M. Westerlund, O. Pletnikova, L. Glauser, L. Yang, Y. Liu, D. A. Swing, M. F. Beal, J. C. Troncoso, J. M. McCaffery, N. A. Jenkins, N. G. Copeland, D. Galter, B. Thomas, M. K. Lee, T. M. Dawson, V. L. Dawson and D. J. Moore (2011). "Dopaminergic neuronal loss, reduced neurite complexity and autophagic abnormalities in transgenic mice expressing G2019S mutant LRRK2." PLoS One **6**(4): e18568.

Reinhardt, P., B. Schmid, L. F. Burbulla, D. C. Schondorf, L. Wagner, M. Glatza, S. Hoing, G. Hargus, S. A. Heck, A. Dhingra, G. Wu, S. Muller, K. Brockmann, T. Kluba, M. Maisel, R. Kruger, D. Berg, Y. Tsytsyura, C. S. Thiel, O. E. Psathaki, J. Klingauf, T. Kuhlmann, M. Klewin, H. Muller, T. Gasser, H. R. Scholer and J. Sterneckert (2013). "Genetic correction of a LRRK2 mutation in human iPSCs links parkinsonian neurodegeneration to ERK-dependent changes in gene expression." Cell Stem Cell **12**(3): 354-367.

Rocca, W. A., S. K. McDonnell, K. J. Strain, J. H. Bower, J. E. Ahlskog, A. Elbaz, D. J. Schaid and D. M. Maraganore (2004). "Familial aggregation of Parkinson's disease: The Mayo Clinic family study." Ann Neurol **56**(4): 495-502.

Roosen, D. A. and M. R. Cookson (2016). "LRRK2 at the interface of autophagosomes, endosomes and lysosomes." Mol Neurodegener **11**(1): 73.

Sanchez-Danes, A., Y. Richaud-Patin, I. Carballo-Carbajal, S. Jimenez-Delgado, C. Caig, S. Mora, C. Di Guglielmo, M. Ezquerra, B. Patel, A. Giralt, J. M. Canals, M. Memo, J. Alberch, J. Lopez-Barneo, M. Vila, A. M. Cuervo, E. Tolosa, A. Consiglio and A. Raya (2012). "Disease-specific phenotypes in dopamine neurons from human iPS-based models of genetic and sporadic Parkinson's disease." EMBO Mol Med **4**(5): 380-395.

Satake, W., Y. Nakabayashi, I. Mizuta, Y. Hirota, C. Ito, M. Kubo, T. Kawaguchi, T. Tsunoda, M. Watanabe, A. Takeda, H. Tomiyama, K. Nakashima, K. Hasegawa, F. Obata, T. Yoshikawa, H. Kawakami, S. Sakoda, M. Yamamoto, N. Hattori, M. Murata, Y. Nakamura and T. Toda (2009). "Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease." Nat Genet **41**(12): 1303-1307.

Seaman, M. N. (2012). "The retromer complex - endosomal protein recycling and beyond." J Cell Sci **125**(Pt 20): 4693-4702.

Sen, S., P. J. Webber and A. B. West (2009). "Dependence of leucine-rich repeat kinase 2 (LRRK2) kinase activity on dimerization." J Biol Chem **284**(52): 36346-36356.

Simon-Sanchez, J., C. Schulte, J. M. Bras, M. Sharma, J. R. Gibbs, D. Berg, C. Paisan-Ruiz, P. Lichtner, S. W. Scholz, D. G. Hernandez, R. Kruger, M. Federoff, C. Klein, A. Goate, J. Perlmutter, M. Bonin, M. A. Nalls, T. Illig, C. Gieger, H. Houlden, M. Steffens, M. S. Okun, B. A. Racette, M. R. Cookson, K. D. Foote, H. H. Fernandez, B. J. Traynor, S. Schreiber, S. Arepalli, R. Zonozi, K. Gwinn, M. van der Brug, G. Lopez, S. J. Chanock, A. Schatzkin, Y. Park, A. Hollenbeck, J. Gao, X. Huang, N. W. Wood, D. Lorenz, G. Deuschl, H. Chen, O. Riess, J. A. Hardy, A. B. Singleton and T. Gasser (2009). "Genome-wide association study reveals genetic risk underlying Parkinson's disease." Nat Genet **41**(12): 1308-1312.

Singleton, A. and J. Hardy (2011). "A generalizable hypothesis for the genetic architecture of disease: pleomorphic risk loci." Hum Mol Genet **20**(R2): R158-162.

Spillantini, M. G., M. L. Schmidt, V. M. Lee, J. Q. Trojanowski, R. Jakes and M. Goedert (1997). "Alpha-synuclein in Lewy bodies." Nature **388**(6645): 839-840.

Steger, M., F. Tonelli, G. Ito, P. Davies, M. Trost, M. Vetter, S. Wachter, E. Lorentzen, G. Duddy, S. Wilson, M. A. Baptista, B. K. Fiske, M. J. Fell, J. A. Morrow, A. D. Reith, D. R. Alessi and M. Mann (2016). "Phosphoproteomics reveals that Parkinson's disease kinase LRRK2 regulates a subset of Rab GTPases." Elife **5**.

Stoessl, A. J., W. W. Martin, M. J. McKeown and V. Sossi (2011). "Advances in imaging in Parkinson's disease." Lancet Neurol **10**(11): 987-1001.

Sveinbjornsdottir, S., A. A. Hicks, T. Jonsson, H. Petursson, G. Gugmundsson, M. L. Frigge, A. Kong, J. R. Gulcher and K. Stefansson (2000). "Familial aggregation of Parkinson's disease in Iceland." N Engl J Med **343**(24): 1765-1770.

Tauchert, M. J., C. Hemonnot, P. Neumann, S. Koster, R. Ficner and A. Dickmanns (2016). "Impact of the crystallization condition on importin-beta conformation." Acta Crystallogr D Struct Biol **72**(Pt 6): 705-717.

Tong, Y., E. Giaime, H. Yamaguchi, T. Ichimura, Y. Liu, H. Si, H. Cai, J. V. Bonventre and J. Shen (2012). "Loss of leucine-rich repeat kinase 2 causes age-dependent bi-phasic alterations of the autophagy pathway." Mol Neurodegener **7**: 2.

Trinh, J., M. Farrer, O. A. Ross and I. Guella (2006). LRRK2-Related Parkinson Disease. GeneReviews((R)). M. P. Adam, H. H. Ardinger, R. A. Pagon et al. Seattle (WA).

van Egmond, W. N. and P. J. van Haastert (2010). "Characterization of the Roco protein family in Dictyostelium discoideum." Eukaryot Cell **9**(5): 751-761.

Varanese, S., Z. Birnbaum, R. Rossi and A. Di Rocco (2011). "Treatment of advanced Parkinson's disease." Parkinsons Dis **2010**: 480260.

Vetter, I. R. and A. Wittinghofer (2001). "The guanine nucleotide-binding switch in three dimensions." Science **294**(5545): 1299-1304.

Wang, T., Y. Liu, X. H. Xu, C. Y. Deng, K. Y. Wu, J. Zhu, X. Q. Fu, M. He and Z. G. Luo (2011). "Lgl1 activation of rab10 promotes axonal membrane trafficking underlying neuronal polarization." Dev Cell **21**(3): 431-444.

Wang, W., I. Perovic, J. Chittuluru, A. Kaganovich, L. T. Nguyen, J. Liao, J. R. Auclair, D. Johnson, A. Landeru, A. K. Simorellis, S. Ju, M. R. Cookson, F. J. Asturias, J. N. Agar, B. N. Webb, C. Kang, D. Ringe, G. A. Petsko, T. C. Pochapsky and Q. Q. Hoang (2011). "A soluble alpha-synuclein construct forms a dynamic tetramer." Proc Natl Acad Sci U S A **108**(43): 17797-17802.

West, A. B., D. J. Moore, S. Biskup, A. Bugayenko, W. W. Smith, C. A. Ross, V. L. Dawson and T. M. Dawson (2005). "Parkinson's disease-associated mutations in leucine-rich repeat kinase 2 augment kinase activity." Proc Natl Acad Sci U S A **102**(46): 16842-16847.

West, A. B., D. J. Moore, C. Choi, S. A. Andrabi, X. Li, D. Dikeman, S. Biskup, Z. Zhang, K. L. Lim, V. L. Dawson and T. M. Dawson (2007). "Parkinson's disease-associated mutations in LRRK2 link enhanced GTP-binding and kinase activities to neuronal toxicity." Hum Mol Genet **16**(2): 223-232.

Winn, M. D., C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson (2011).

"Overview of the CCP4 suite and current developments." Acta Crystallogr D Biol Crystallogr **67**(Pt 4): 235-242.

Winn, M. D., G. N. Murshudov and M. Z. Papiz (2003). "Macromolecular TLS refinement in REFMAC at moderate resolutions." Methods Enzymol **374**: 300-321.

Wittinghofer, A. and E. F. Pai (1991). "The structure of Ras protein: a model for a universal molecular switch." Trends Biochem Sci **16**(10): 382-387.

Xing, Y., T. Bocking, M. Wolf, N. Grigorieff, T. Kirchhausen and S. C. Harrison (2010). "Structure of clathrin coat with bound Hsc70 and auxilin: mechanism of Hsc70-facilitated disassembly." EMBO J **29**(3): 655-665.

Xu, C. and J. Min (2011). "Structure and function of WD40 domain proteins." Protein Cell **2**(3): 202-214.

Zabetian, C. P., A. Samii, A. D. Mosley, J. W. Roberts, B. C. Leis, D. Yearout, W. H. Raskind and A. Griffith (2005). "A clinic-based study of the LRRK2 gene in Parkinson disease yields new mutations." Neurology **65**(5): 741-744.

Zhang, C. X., A. E. Engqvist-Goldstein, S. Carreno, D. J. Owen, E. Smythe and D. G. Drubin (2005). "Multiple roles for cyclin G-associated kinase in clathrin-mediated sorting events." Traffic **6**(12): 1103-1113.

Zimprich, A., S. Biskup, P. Leitner, P. Lichtner, M. Farrer, S. Lincoln, J. Kachergus, M. Hulihan, R. J. Uitti, D. B. Calne, A. J. Stoessl, R. F. Pfeiffer, N. Patenge, I. C. Carbajal, P. Vieregge, F. Asmus, B. Muller-Myhsok, D. W. Dickson, T. Meitinger, T. M. Strom, Z. K. Wszolek and T. Gasser (2004). "Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology." Neuron **44**(4): 601-607.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46(W1), W296-W303 (2018).

Bienert, S., Waterhouse, A., de Beer, T.A.P., Tauriello, G., Studer, G., Bordoli, L., Schwede, T. The SWISS-MODEL Repository - new features and functionality. Nucleic Acids Res. 45, D313-D319 (2017).

Guex, N., Peitsch, M.C., Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis 30, S162-S173 (2009).

Benkert, P., Biasini, M., Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27, 343-350 (2011).

Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L., Schwede, T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. Scientific Reports 7 (2017).

CURRICULUM VITAE

Chunxiang Wu

Education:

2011-2019 Ph. D. in Biochemistry and Molecular Biology,

Indiana University, Indianapolis, Indiana, United States

Advisor: Dr. Quyen Q. Hoang

Committee: Dr. Tatiana M. Foroud

Dr. Thomas D. Hurley

Dr. Steven M. Johnson

Dr. Zhong-Yin Zhang

2007-2011 B.S. in Biological Science,

Anhui Agriculture University, Hefei, Anhui, China

Research Experience:

2012-2019 Research Assistant

Department of Biochemistry and Molecular Biology, Indiana University

Advisor: Dr. Quyen Q. Hoang

Projects: Structure and function of Parkinson's disease associated protein, LRRK2

Publications:

2019 **Chun-Xiang Wu**, Jingling Liao, Yangshin Park, Xylena Reed, Victoria A. Engel, Neo C. Hoang, Yuichiro Takagi, Steven M. Johnson, Mu Wang, Mark Federici, Jeremy Nichols, Ruslan Sanishvili, Mark R. Cookson, and Quyen Q. Hoang. (2019) Parkinson's disease-associated mutations in the GTPase domain of LRRK2 impair its nucleotide dependent conformational dynamics. J Biol Chem. 2019 Feb 24 [Epub ahead of print]

Chun-Xiang Wu, Jingling Liao, Yangshin Park, Neo C. Hoang, Victoria A. Engel, Ruslan Sanishvili, Yuichiro Takagi, Steven M. Johnson, Mu Wang, Mark Federici, Jeremy Nichols, Mark R. Cookson, and Quyen Q. Hoang. (2019) Structural basis for conformational plasticity of the Parkinson's disease-associated protein LRRK2. (Manuscript in preparation for PNAS)

2018 Xiaorong Huang, **Chunxiang Wu**, Yangshin Park, Xuwei Long, Quyen Q. Hoang, and Jingling Liao. (2018) The Parkinson's disease–associated mutation N1437H impairs conformational dynamics in the G domain of LRRK2. FASEB J. 2018 Dec 28 [Epub ahead of print]

2016 Wei Wang, Linh T. T. Nguyen, Christopher Burlak, Fariba Chegini, Feng Guo, Tim Chataway, Shulin Ju, Oriana S. Fisher, David W. Miller, Debajyoti Datta, Fang Wu, **Chun-Xiang Wu**, Anuradha Landeru, James A. Wells, Mark R. Cookson, Matthew B. Boxer, Craig J. Thomas, Wei Ping Gai, Dagmar Ringe, Gregory A. Petsko, and Quyen Q. Hoang. (2016)

Caspase-1 causes truncation and aggregation of the Parkinson's disease-associated protein α-synuclein. PNAS August 23, 2016 113 (34) 9587-9592

2015 Romila Mascarenhas, Pei W. Thomas, **Chun-Xiang Wu**, Boguslaw P. Nocek, Quyen Q. Hoang, Dali Liu, and Walter Fast. Structural and Biochemical Characterization of AidC, a Quorum-Quenching Lactonase with Atypical Selectivity. Biochemistry, 2015, 54 (28), pp 4342–4353

2014 Jingling Liao, **Chun-Xiang Wu**, Christopher Burlak, Sheng Zhang, Heather Sahm, Mu Wang, Zhong-Yin Zhang, Kurt W. Vogel, Mark Federici, Steve M. Riddle, R. Jeremy Nichols, Dali Liu, Mark R. Cookson, Todd A. Stone, and Quyen Q. Hoang. PNAS March 18, 2014 111 (11) 4055-4060

Shinya Tashiro, Jose M. M. Caaveiro, **Chun-Xiang Wu**, Quyen Q. Hoang, and Kouhei Tsumoto. Thermodynamic and Structural Characterization of the Specific Binding of Zn (II) to Human Protein DJ-1. Biochemistry, 2014, 53 (14), pp 2218–2220

Posters at Conferences:

2017 Gordon Research Conference:

Parkinson's disease

Poster Title: Biochemical Bases for the pathogenic effects of R1441 mutations in the ROC domain of LRRK2

Authors: **Chunxiang Wu**, Jingling Liao, Yangshin Park, Alexe' Engel, and Quyen Hoang

2014 Keystone Symposia:

Parkinson's disease: Genetics, Mechanisms and Therapeutics

Poster Title: The Parkinson's disease-associated mutations R1441H in LRRK2 prolongs the 'active-state' of its GTPase domain

Authors: Jingling Liao, **Chunxiang Wu**, Christopher Burlak, Sheng Zhang, Heather Sahm, Mu Wang, Zhong-Yin Zhang, Kurt W. Vogel, Steve Riddle, Mark Federici, Jeremy Nichols, Dali Liu, Mark R. Cookson, Todd A. Stone, and Quyen Q. Hoang.